UNDERSTANDING THE IMPACT OF INTERNAL VARIABILITY ON ESTIMATES OF THE TRANSIENT CLIMATE RESPONSE

A Thesis

by

BROOKE K. ADAMS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Andrew Dessler
Committee Members,	Yangyang Xu
	Gerald North
Head of Department,	Ping Yang

August 2018

Major Subject: Atmospheric Sciences

ABSTRACT

The transient climate response (TCR) is useful for quantifying the Earth's short-term (decadal to century timescales) response to an increase in greenhouse gases. The TCR cannot be directly observed, so it is often estimated using the historical temperature record. One important question is the extent to which internal variability confounds these estimates of TCR. To quantify the uncertainty caused by internal variability, we analyze output from a 100-member ensemble of a fully coupled global climate model (Max Planck Institute Earth Science Model, MPI-ESM1.1). Each of the 100 ensemble members has identical historical forcing and model physics but unique initial conditions. The model members consequently evolve differently in time. The TCR estimated from the historical ensemble yields values ranging from 1.34 to 1.9 K, where the spread can be attributed to internal variability alone. We can understand this range using a framework derived from the Earth's energy balance. A majority of the variability observed in the ensemble estimates of TCR is due to heat transport into the deep oceans, ensemble members that transport more heat to the deep ocean have lower TCRs. Another key factor, is the amount of forcing that is radiated back to space. Together, these two factors explain 71% of the variance in the TCR estimates from the large ensemble. Due to the demonstrated effects of internal variability, an estimate of the TCR from the historical record (a single realization out of many possible climate histories) could deviate significantly from the climate system's true value.

DEDICATION

This thesis is dedicated to my Grandma Adams. Thank you for your love and support, teaching me the value of education, and for all of the cookies.

ACKNOWLEDGMENTS

I would like to thank Dr. Dessler, my advisor, for all his time, effort, and patience. I am appreciative of the guidance he has given me and will continue to apply it throughout my life. Thank you to Dr. Xu and Dr. North, my committee members, for critiquing my work and for being excellent teachers. I would also like to thank Steve Schroeder for providing me feedback on my thesis.

I would also like to thank my group members and friends here at Texas A&M for supporting me and answering all my questions along the way. Thanks for being great colleagues and friends.

I would like to express my gratitude for the love and support from my parents, grandparents, and my sisters. Thank you for motivating and inspiring me throughout my academic career. Finally, thank you Jordan Gregg for encouraging me to be the best I can be.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a thesis committee consisting of Dr. Dessler and Dr. Xu of the Department of Atmospheric Sciences and Dr. North of the Department of Oceanography. Graduate study was supported by a National Science Foundation grant AGS-1661861 to Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

1.1 Earth's energy balance

Following the first law of thermodynamics, energy in the Earth system must be conserved. Thus, when a top-of-atmosphere (TOA) energy imbalance (hereafter referred to as forcing, ΔF) is imposed on the planet, such as from an increase in carbon dioxide, the planet responds by warming or cooling. This process is traditionally described by the linearized energy balance equation [e.g. Dessler and Zelinka, 2015]:

$$\Delta R = \Delta F + \lambda \Delta T \tag{1.1}$$

where ΔT is the change in global average surface temperature, and λ is the change in TOA flux per unit change in ΔT. Changes in ΔT can be considered the response of the system to the forcing. ΔR is the resulting TOA flux imbalance from the combined forcing and response.

While Eq. 1.1 is frequently said to describe "energy balance," it actually balances power. We can convert this to a true energy balance by integrating each term with respect to time and over the surface area of the earth:

$$\int \Delta R dt = \int \Delta F dt + \int \lambda \Delta T dt \tag{1.2}$$

This provides a physical framework for understanding climate change. $\int \Delta$F dt is the total energy (in Joules) trapped by the imposed forcing. This energy is either stored in the Earth system, equal to $\int \Delta$R dt, or is radiated to space, equal to $\int \lambda \Delta$T dt. Previous investigators have taken this energy-budget approach and estimate that about 2/3 of forced energy is radiated back to space and about 1/3 is stored as accumulated energy [Murphy et al., 2009, Huber and Knutti, 2012]. It is also found that about 90% of the energy stored in the Earth ($\int \Delta$R dt) is sequestered into ocean heat uptake [Von Schuckmann et al., 2016, Allan, 2017].

1.2 Climate sensitivity

Climate sensitivity is an estimate of the amount of warming that occurs in response to a fixed forcing. The most commonly used metric is the Equilibrium Climate Sensitivity (ECS), defined to be the equilibrium temperature response to a doubling of carbon dioxide. We can rearrange Eq. 1.1 to yield an equation for ECS:

$$ECS = \frac{-F_{2xCO2}}{\lambda} \tag{1.3}$$

where F_{2xCO2} is the forcing from doubled CO_2.

Because of the thermal inertia of the oceans, it takes millennia for the Earth to reach equilibrium after the imposition of a forcing. This has led some researchers to focus on a different metric, the transient climate response (TCR). It is defined as the change in global mean surface temperature in response to a 1% /year increase in CO_2 at the time of doubling, 70 years. Compared to the ECS, the TCR characterizes the Earth's response to a doubling of CO_2 on shorter, more policy relevant time scales.

1.2.1 Estimates of the TCR from the historical record

Despite the fact that the Earth has not been forced by a simple 1% per year increase in CO_2, we can estimate TCR from the observational record with the equation [Gregory et al., 2002; Otto et al., 2013; Lewis and Curry, 2015]:

$$TCR = \Delta T * \frac{F_{2xCO2}}{\Delta F} \tag{1.4}$$

where ΔT is the temperature change (relative to some approximately pre-industrial base period), ΔF is historical forcing (relative to the same period), and F_{2xCO2} is the forcing from doubled CO_2. The choice of the reference and base periods used in the delta terms is discussed in section 2.

Several studies have shown that estimates of the TCR made from historical observations using

Eq. 1.4 fall in the lower range of the estimates of TCR from GCMs [Otto et al., 2013; Fyfe et al., 2013; Mauritsen and Pincus, 2017; Medhaug et al., 2017; Lewis and Curry, 2015; Armour, 2017]. This discrepancy between model and observational estimates could be explained several ways. One possibility, of course, is that the models are incorrect and overestimate the warming from CO_2 [Stott et al., 2013; Fyfe et al., 2013]. However, there are also issues with the observational estimates of TCR. Spatial inhomogeneities in the surface temperature record have been shown to affect estimates of the TCR [Richardson et al., 2016; Cowtan and Way, 2014]. The magnitude of external radiative forcing is also quite uncertain, specifically the forcing from aerosols [Fyfe et al., 2013; Lewis and Curry, 2015; Forster, 2016].

Recently, it has become apparent that internal variability of the climate system can confound estimates of climate sensitivity [Frankcombe et al., 2015; Olson et al., 2013; Huber et al., 2014]. Several analyses have shown that ECS estimates from the historical record are strongly affected by internal variability in the pattern of surface warming [Zhou et al., 2016; Gregory and Andrews, 2016; Andrews and Webb, 2018, Dessler et al., 2018]. The implication of this is that ECS inferred from the 20^{th} century is going to be low biased [Marvel et al., 2018].

TCR, on the other hand, should be less impacted by the pattern of surface warming, but more impacted by variability in the total amount of surface warming (Eq. 1.4) although the total amount of warming may be linked to the pattern [Brown et al., 2016]. Previous studies have shown that internal variability hinders the predictability of the total amount of warming (ΔT) [Deser et al., 2012; Frankcombe et al., 2015; Hawkins and Sutton, 2009; Marotzke and Forster, 2015; Skeie et al., 2014; Padilla et al., 2011]. Since observational estimates of the TCR involve scaling ΔT by a constant fraction of $F_{2xCO2}/\Delta F$ (Eq. 1.4), the effect of internal variability on the total amount of warming translates into uncertainty in the TCR [Gregory and Forster, 2008; Lewis and Curry, 2015]. Huber et al. (2014) quantified the effect of internal variability on estimates of the TCR and found the estimates to vary on the order of 0.1 K (+/- 1 σ) from the ensemble mean of about 1.8 K within a 20-member initial condition experiment of the Bern2.5D model.

1.2.2 Quantifying internal variability

Model ensembles provide a tool to study the impact of internal variability on estimates of climate sensitivity. The most useful ensembles contain many runs of a single model with identical physics and external forcing but different initial conditions. As each member evolves differently in time, the climate predictions of the model members spread out in response to different realizations of internal climate variability [Kay et al., 2015; Hedemann et al., 2017]. In fact, one can think of the historical record as one member of a theoretical ensemble of Earth's climate trajectories since pre-industrial. Thus, the model ensemble gives us an estimate of what the distribution of alternative climate histories may have looked like.

Dessler et al. (2018) analyzed a 100-member ensemble of the Max Planck Institute Earth System Model (MPI-ESM1.1) to characterize the impact of internal variability on estimates of the ECS; they found that internal variability can lead to large errors in the inferred ECS. ECS estimates from the large ensemble were found to range from 2.1 to 3.9 K [Dessler et al., 2018]. Observations are based on a single realization of internal climate variability. A large spread in the distribution of climate sensitivity, due to internal variability, can lead to significant differences between estimates from the historical record and the climate system's true value of climate sensitivity [Olson et al., 2013; Marvel et al., 2018; Dessler et al., 2018].

1.3 Thesis goals

The goal of this thesis is to quantify the impact of internal variability on estimates of the TCR from the 20^{th} century historical record. The first section provides an introduction to the TCR. The second section describes the model output used and the methods of calculating the variables used in analysis. The third section quantifies the spread of estimates of the TCR made from the 20^{th} century record and investigates the spatial pattern of variability. The fourth section uses an energy balance framework to understand what is controlling the variability in the TCR. The fifth section provides a description of the TCR from the 1^{st} and the 2^{nd} doubling of CO_2. The sixth section provides a summary of our work.

2. MODEL DESCRIPTION & METHODS

2.1 Model data

In our analysis, we use output from the Max Planck Institute Earth System Model version 1.1 (MPI-ESM1.1) (referred to as the large ensemble). MPI-ESM1.1 is a coupled atmosphere-ocean climate model comprised of ECHAM6.3 for the land and atmosphere and Max Planck Institute-Ocean Model (MPI-OM) for the ocean. The atmospheric model has T63 spectral truncation, or 1.9 degrees horizontal resolution, with 47 vertical levels. The ocean model has about 1.5 degrees horizontal resolution and 40 vertical levels [Dessler et al., 2018]. The large ensemble contains 100 members that simulate the years 1850-2005 (Fig. 2.2) and use identical evolutions of historical natural and anthropogenic forcing. The 100 simulations only differ in their initial conditions, derived from varying points in a 2000-year control run. As the model members evolve differently in time, their spread can be used to estimate the distribution of internal variability in the model. Further details of the model are described in Giorgetta et al. (2013), Dessler et al. (2018), and Hedemann et al. (2017).

2.2 Choice of base period

Otto et al. (2013) and Lewis and Curry (2015) have found that equation 1.4 is a valid estimator of the TCR since the forcing in the last 70 years has increased at an approximately linear rate. All delta terms in equation 1.1 are relative to a base period, which can have a large effect on estimates of the TCR [Lewis and Curry, 2015]. According to Lewis and Curry, the base period should not be later than 1950 and the final or reference period must be recent. Three base periods are investigated here: Preindustrial control, 1859-1882 average, and 1850-1860 average. Both the preindustrial control and 1859-1882 average base periods yield an average TCR estimate close to the true TCR of the system (1.81 K, see section 2.5). The TCR estimated using the 1850-1860 base period yields a low biased estimate, because of the state of the forcing during that time period. Previous studies use the 1859-1882 base period because it has little influence from volcanic

eruptions [Mauritsen and Pincus, 2017; Lewis and Curry, 2015]. This is ideal because in addition to forcing values closer to present day values, feedbacks due to a volcanic eruption are physically different than those due to greenhouse gases [Forster, 2016; Lewis and Curry, 2015]. For these reasons, the 1859-1882 base period is selected for all historical analyses in this study. For the 1% per year forced simulation, Δ terms are investigated relative to the 1850 average.

2.3 Variables used for analysis

2.3.1 Forcing and TOA flux in the large ensemble

All 100 ensemble members are forced with the CMIP5 historical forcing time series from 1850-2005. Giorgetta et al. (2013) describes the details of MPI-ESM. The large ensemble is forced by CMIP5 historical forcing which includes both natural forcing and anthropogenic forcing. The natural forcings are from Earth's orbit, sun spots, tropospheric aerosols, and stratospheric aerosols from volcanic eruptions. Well mixed greenhouse gases, carbon dioxide, methane, nitrous oxide, chlorofluorocarbons (CFC-11 and CFC-12), and ozone contribute to the anthropogenic forcing. The historical forcing also simulates land use change, which effects the fraction of incoming solar radiation that is reflected back to space (planetary albedo) [Giorgetta et al., 2013].

Figure 2.1 is a time series of the forcing relative to the 1859-1882 average (so the runs start with slightly negative forcing in the 1850s). The trend displays forced responses to volcanic eruptions, which are represented by large negative anomalies in the forcing, including Krakatoa in 1883, Mount Agung in 1963, and Pinatubo in 1991. There is also a general increase in the forcing from about -0.5 to 2.3 Wm^{-2} throughout the length of the run due to an increase in anthropogenic greenhouse gases.

The CO_2 increase is also shown in the top of the atmosphere energy imbalance (ΔR, relative to 1859-1882 average). A positive imbalance represents more energy stored in the earth- atmosphere system (including the oceans) than re-radiating out, which occurs in the presence of greenhouse gases. Negative ΔR values occur during volcanic eruptions. The mean TOA imbalance at the beginning of the run is about -0.1 Wm^{-2} and by 2005 the ensemble average of ΔR is about 0.7

Wm^{-2}.

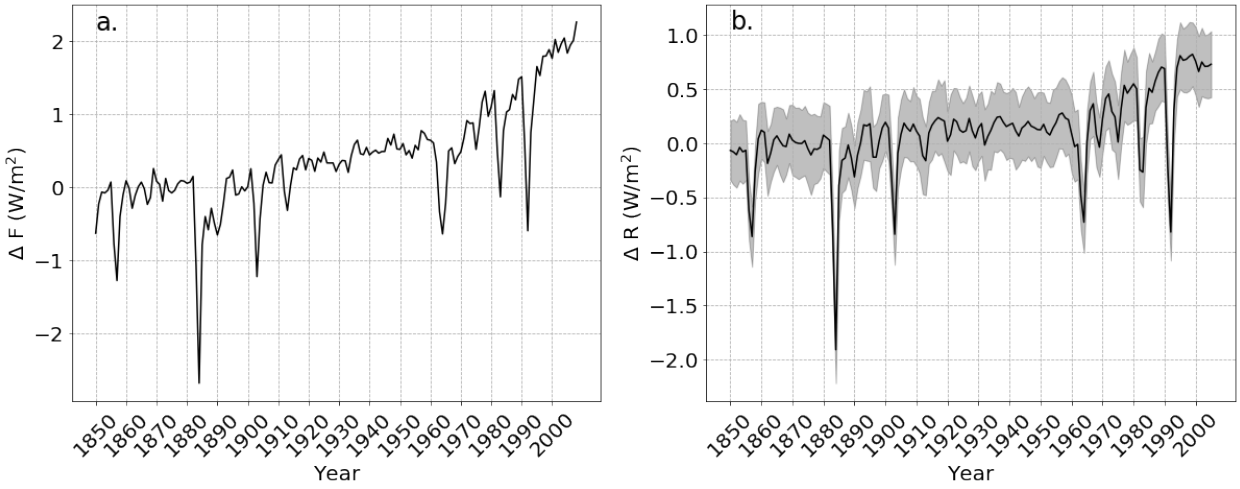

Figure 2.1: Panel a: Time series of ΔF. Panel b: ΔR for the historical ensemble, relative to the 1859-1882 average. The grey shading represents the ensemble mean ± one standard deviation.

2.3.2 Historical ΔT

The time series of ΔT (1995-2005 average relative to the 1859-1882 average) of all 100 model members is shown in figure 2.2. The global average quantities in this study are done with a cosine weighting, to ensure that each grid box contributes in proportion to its area. The forced responses within the ensemble are represented by periods of time where all the 100 model members move in sync in figure 2.2. There are clear forced responses due to volcanic eruptions, i.e. in 1883 and 1963. The overall upward trend present throughout the length of the run is also representative of a forced response in reaction to an increase in greenhouse gases; the ensemble average of temperature increases to 0.9 K by the end of the run. In contrast to the forced responses, the spread of individual ensemble members about the ensemble mean (white line) is representative of internal variability or an unforced climate response. There is considerable spread in the ensemble temperature at the end of the run due to internal variability (Fig. 2.2). ΔT at year 2005 ranges from 0.6 to 1.3 K. Because TCR is proportional to ΔT (Eq. 1.4), the 0.7 K spread translates directly into

uncertainty in estimates of the TCR (discussed in section 3). Historical (1850-2005, relative to the base period) and $2xCO_2$ forcing, calculated from fixed sea-surface temperature runs, are 1.83 Wm^{-2} and 3.7 Wm^{-2}, respectively, so the ratio $F_{2xCO2}/\Delta F$ in Eq. 1.4 is a constant (2.02).

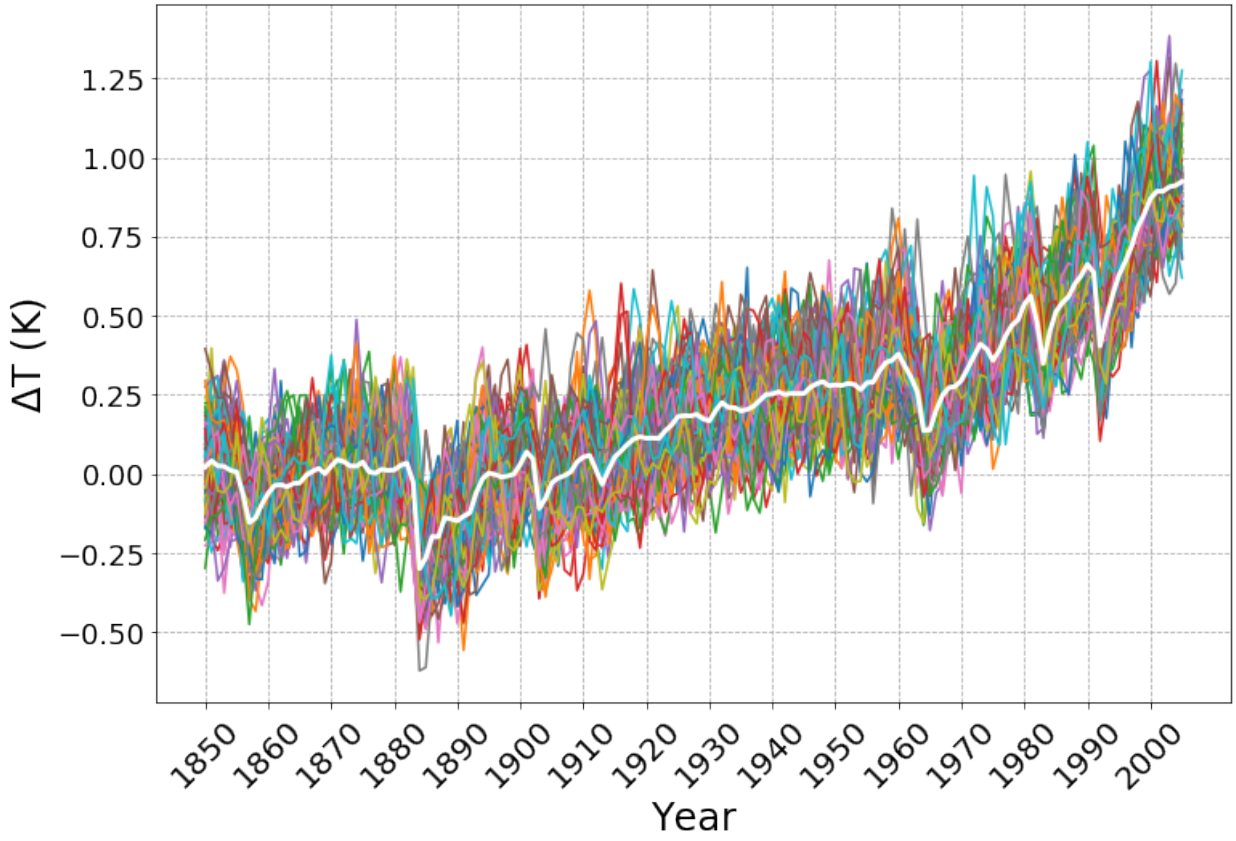

Figure 2.2: Adapted from Dessler et al. (2018). Time series of change in temperature (relative to 1859-1882 average) from 1850-2005. The colored lines represent each individual model member and the white line is the mean of all 100 of the ensemble members.

2.3.3 λ based on historical forcing ensemble

The climate feedback parameter, lambda (λ), has units of Wm^{-2} K^{-1} and measures the rate at which energy is radiated out to space. λ encompasses several different feedbacks which act to amplify or depress the initial magnitude of warming. Climate feedbacks are therefore an important controlling factor in the magnitude of Earth's surface temperature increase in response to a positive

external forcing. The overall sign of the climate feedback parameter on earth is negative, thus in equation 1.1 the $\lambda * \Delta$T counteracts the external forcing term (ΔF).

λ is obtained for each ensemble member using a rearranged version of equation 1.1:

$$\lambda = \frac{(\Delta R - \Delta F)}{\Delta T} \tag{2.1}$$

The term R in equation 2.1 represents the top of the atmosphere (TOA) radiative imbalance (Fig 2.1b). This value is obtained by adding the short wave and the long wave radiation at the TOA. The difference of the TOA flux and the forcing (Fig 2.1) is divided by the change in surface temperature for each ensemble member to derive λ.

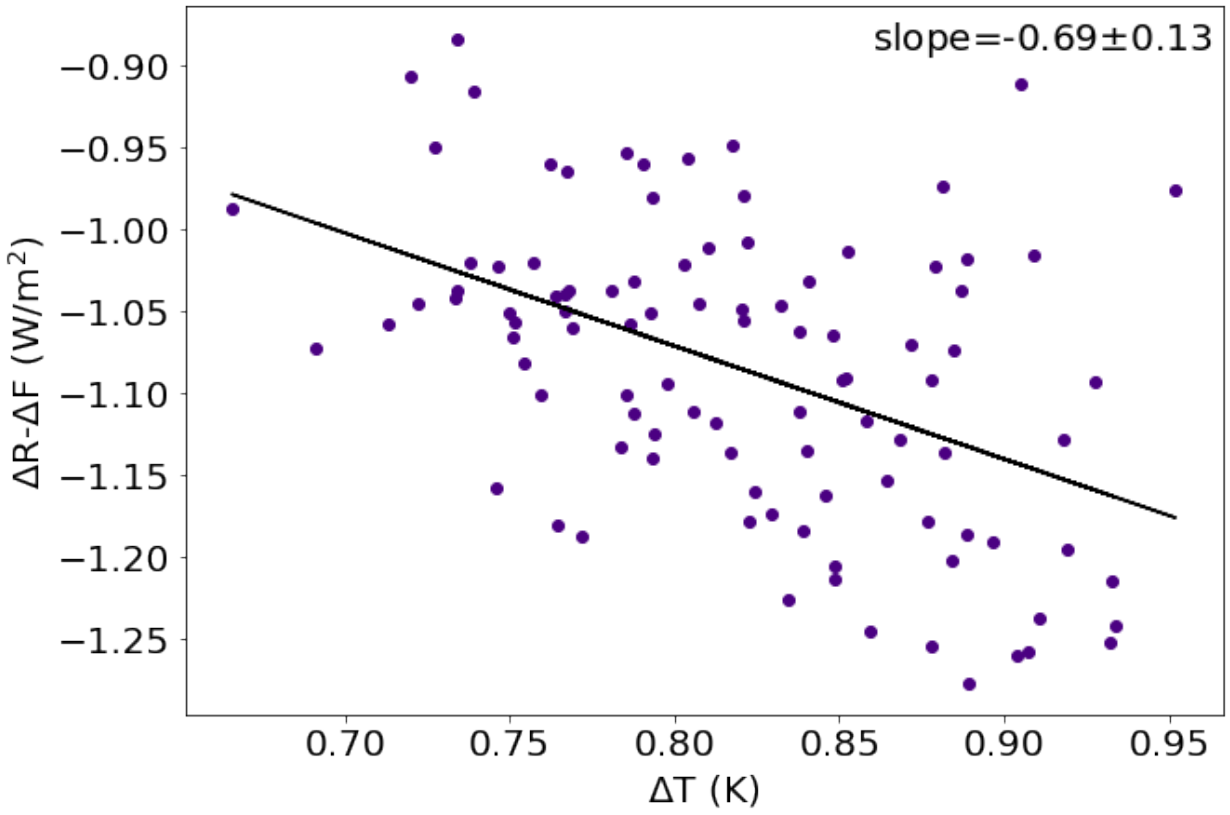

Figure 2.3: Regression of ΔR-ΔF vs. ΔT. The slope of the regression is equal to λ .

In figure 2.3, ΔR-ΔF is plotted versus ΔT for all 100 ensemble members (Eq. 2.1). Following equation 2.1, the slope of the best fit line is equal to the ensemble average of λ . The linear regression yields -0.69 $\pm$ 0.13 Wm^{-2} K^{-1} as the ensemble average value of λ . The negative value of λ describes the fact that as the earth heats up, it radiates more energy to space, as described by the Stefan Boltzmann law. However, this regression shows that λ is rather uncertain in the large ensemble. There is considerable spread in ΔR-ΔF for each value of ΔT (Fig. 2.3) [Dessler et al., 2018]. The relationship of λ to the TCR will be investigated in subsequent sections.

2.4 ΔT with 1% per year CO_2 forcing scenario

The MPI-ESM2M1.1 model has an ensemble of 68 model members that are forced by a scenario with a compounded 1% per year increase in CO_2. The forcing (from CO_2) is increasing at 1% per year and has a relatively constant slope throughout the length of the run (Fig. 2.4). Year 70 represents the time of doubling of CO_2 and is about equal to 3.7 Wm^{-2} (dashed red line). The ensemble average of the TCR, calculated at the time of CO_2 doubling, is 1.81 K. The second half of the run (to year 140) is discussed in section 5. This forcing scenario solely increases the forcing from CO_2 and is initialized from the MPI-ESM preindustrial control run [Giorgetta et al., 2013]. For example, other WMGHG are fixed at their respective concentrations from year 1850. Giorgetta et al. (2013) provide a detailed description of the state of the climate in the preindustrial control run.

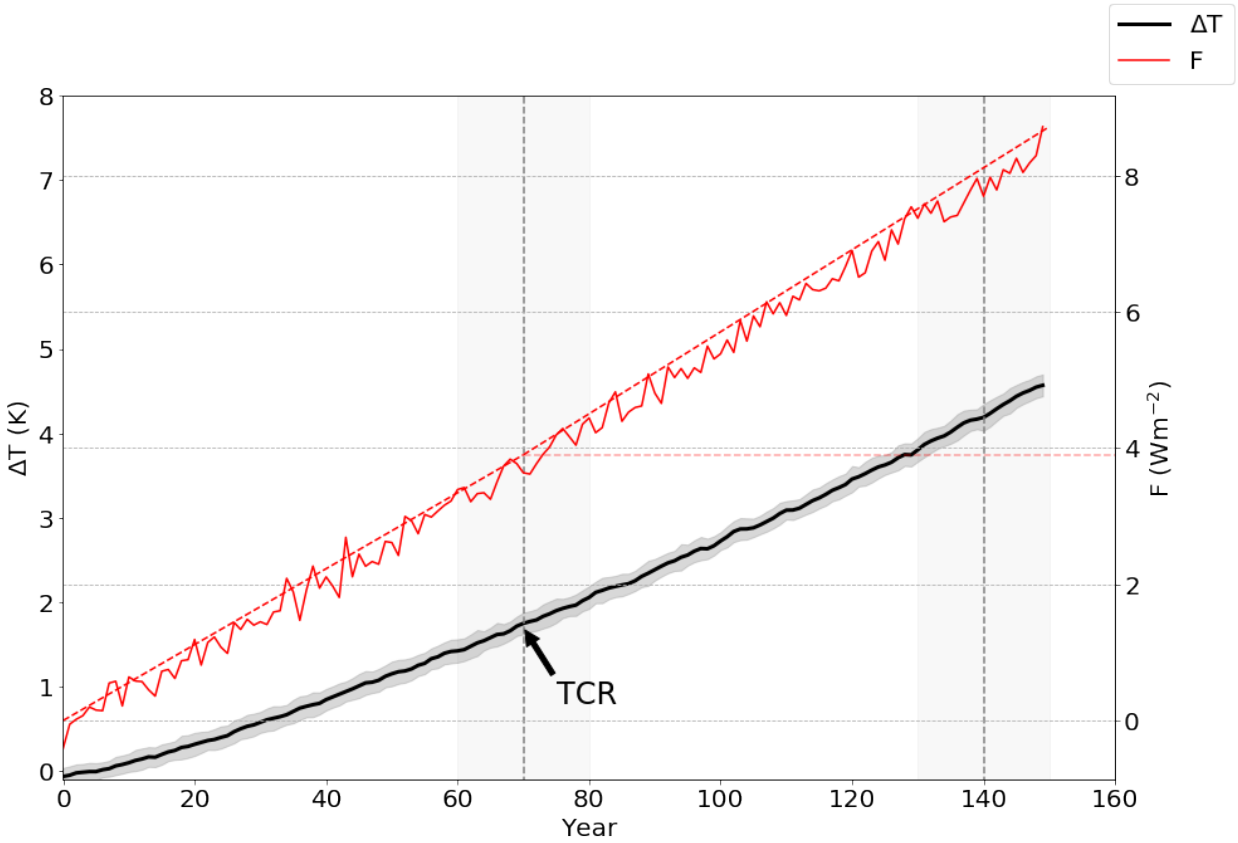

Figure 2.4: Forcing and ΔT time series for the 1% per year CO_2 forcing scenario. The grey shading around ΔT represents the ensemble mean plus and minus one standard deviation. The forcing data is the same for all 68 model members, shown in the solid red line. The dashed red line represents the constant increase of CO_2 to 3.9 Wm^{-2}, the forcing for doubled CO_2. The definition of the TCR, or the temperature increase at the time of CO_2 doubling, is annotated at year 70.

3. THE EFFECT OF INTERNAL VARIABILITY ON ESTIMATES OF THE TCR

The TCR from each ensemble member has been estimated using Eq. 1.4; the resulting distribution of TCR is shown in Fig. 3.1. The 5-95% range of the ensemble's TCR varies from 1.46 K to 1.86 K (dashed lines), and the total range varies from 1.34 K to 1.92 K. This spread in TCR can be attributed to internal variability and we conclude from it that any single estimate of TCR from the historical record may not be indicative of the climate system's actual TCR.

The MPI-ESM1.1's true TCR (derived by averaging the warming in year 70 from a 68-member ensemble of the model forced with a 1% per year increase in CO_2) is 1.81 K, the 5-95% confidence interval varies from 1.59 K to 2.09 K (Fig. 3.1, blue line). This value is, on average, 0.15 K higher than the mean TCR estimated from model driven by historical forcing using the approximation in equation 1.4. Previous studies have noted that observational estimates of TCR tend to be lower than model estimates from the 1% per year forcing scenario [Mauritsen and Pincus, 2017; Lewis and Curry, 2015]. Thus, estimates of the TCR made from the historical record in the large ensemble are both inaccurate and imprecise. The observed spread in estimates of the TCR is evidence that internal variability can add uncertainty to estimates of the TCR made from the historical record.

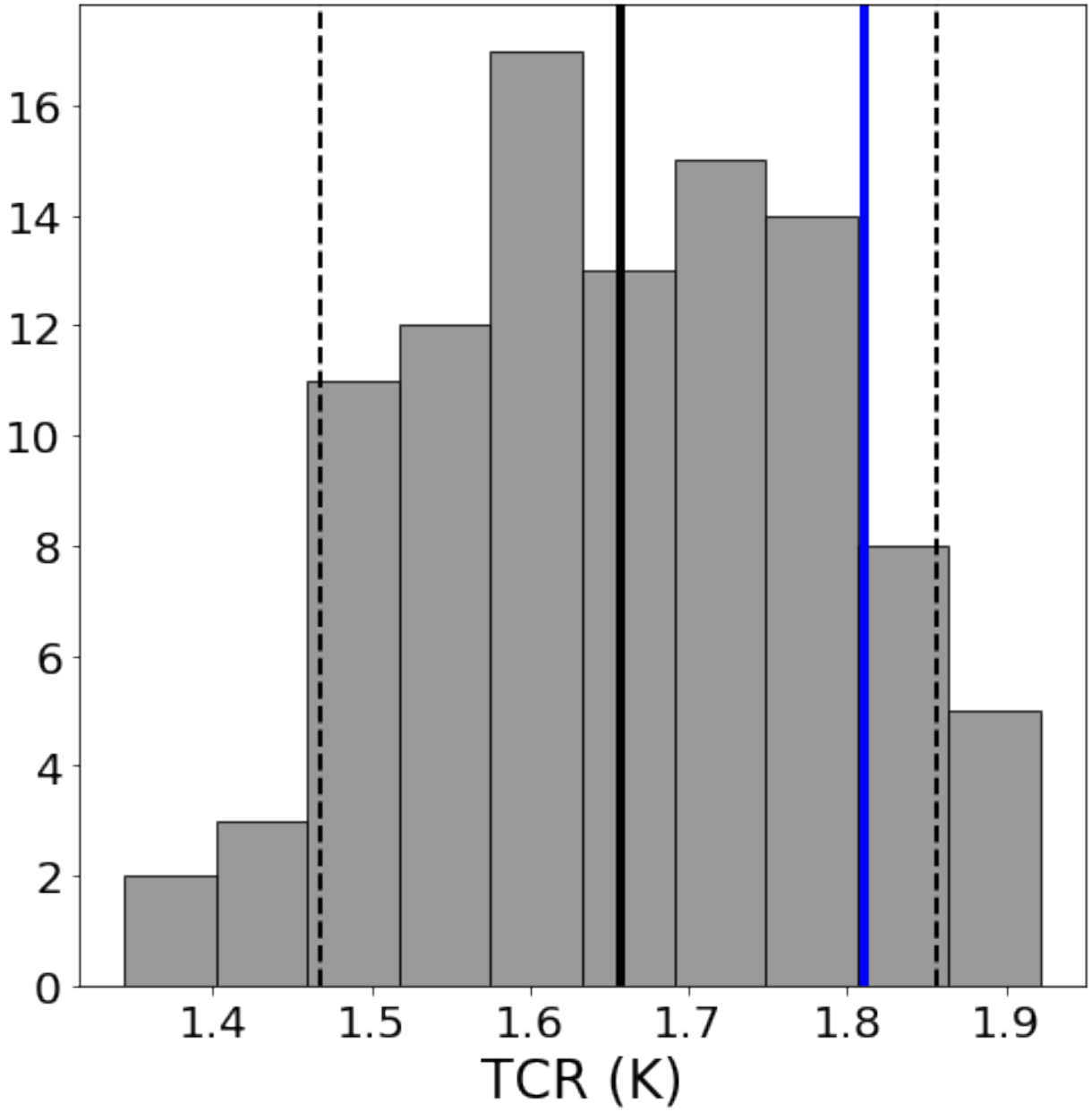

Figure 3.1: PDF of historically estimated TCR using equation 1.4. The solid black line represents the average TCR (50% percentile), and the dashed lines are the 5th and 95th percentile. The solid grey line represents the true TCR of the model. This is calculated by observing the average temperature change at year 70 in a 68-member ensemble of 1% /yr. increase in CO_2 forcing scenario.

3.1 Spatial pattern of ΔT

ΔT vs. the sine of latitude is plotted in figure 3.2 for each model member. The sine of latitude is used as the scale of the x-axis to account for the decrease in grid box size with an increased distance away from the equator. As in figure2.1, the spread around the model mean (white line) represents internal variability. The largest variability is observed in the polar regions (Fig. 3.2). In the Antarctic region, ΔT ranges from about 1.6 K to 0.25 K. In the Arctic region the ensemble average warming is maximized at about 2.75 K, the ensemble also displays a large spread about this mean, varying from 1.5 K to 4.0 K. The large variability and maximum of warming in the Arctic region is likely a consequence of the ice-albedo feedback [Screen and Simmonds, 2010]. The standard deviation of the ensemble from 60°S-30°N remains below 0.75 K. There is a minimum spread in the ensemble members in the Southern Ocean region (60°S) (Fig. 3.2). The Southern Ocean has a very high heat capacity; thus, it is realizing the least amount of surface warming. The tropics also have a relatively high heat capacity which acts as a moderator for its temperature change.

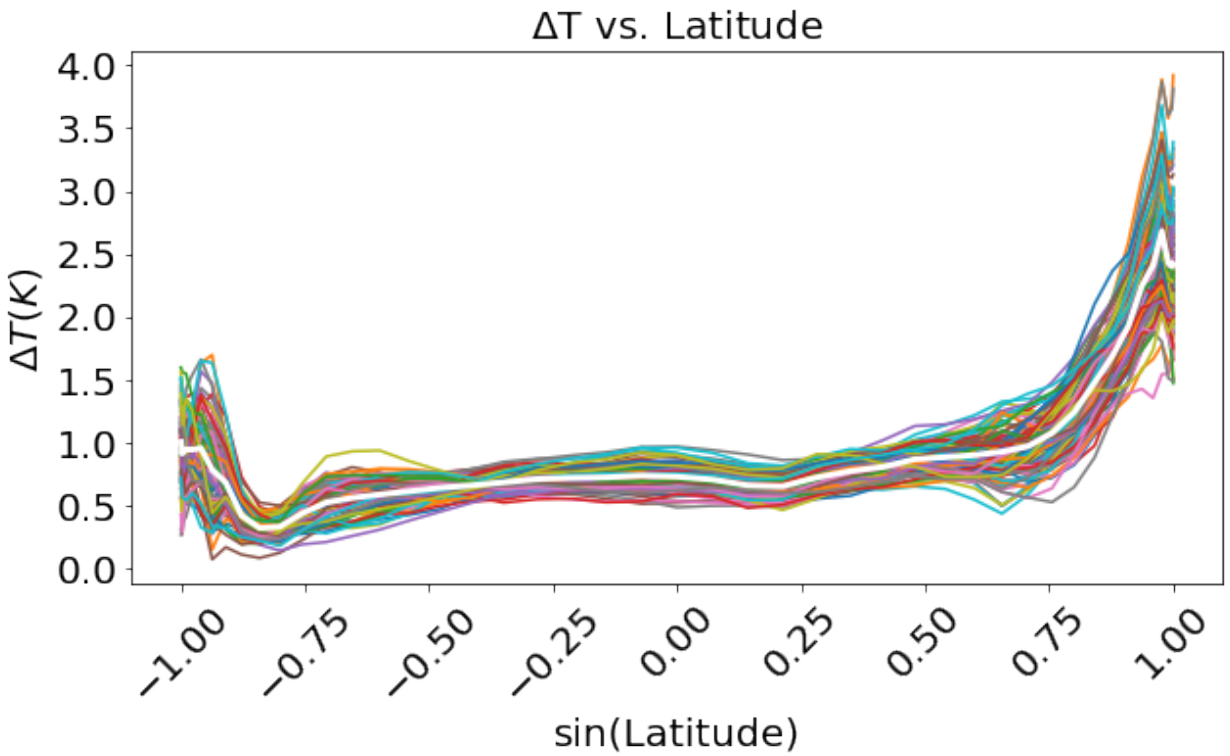

Figure 3.2: Ensemble of ΔT vs. sine of latitude for all ensemble members. The white line represents the model mean, the colored lines represent each individual ensemble member. ΔT is calculated as the 1995-2005 average with reference to the 1859-1882 base-period.

3.2 Spatial pattern of TCR

A global view of the TCR computed from the historically forced ensemble is shown in figure 3.3a. As expected, the TCR has the largest magnitude of warming in the Northern Hemisphere (Fig. 3.3a). Throughout this region, warming ranges from 5-6 K, and a maximum occurs in the northern region of the Barents Sea, with about 10 K of warming (Fig. 3.3a). A local minimum occurs in the Atlantic Ocean south of Greenland. In the northern midlatitudes, the warming is about 2 K over the ocean and is slightly greater over land (2 K- 4 K). A minimum TCR (between 0 and 1 K) exists in the Southern Ocean region.

The standard deviations of the 100-member historically forced ensemble provide a spatial representation of the variability in the data (Fig. 3.3b). There is a near zero standard deviation throughout the tropics and subtropical regions in the northern and southern hemispheres. The regions of

maximum standard deviation are in the Antarctic and Arctic regions, with about 3 K spread between ensemble members (Fig. 3.3b). There are other local maxima in standard deviations located over Alaska (sigma= 1.8 K), off the Eastern coast of Greenland (sigma= 2.1 K), and the East coast of Japan (sigma= 2.3 K). Thus, a majority of the internal variability occurs in the high latitude regions, with some contribution from the northern hemisphere midlatitudes.

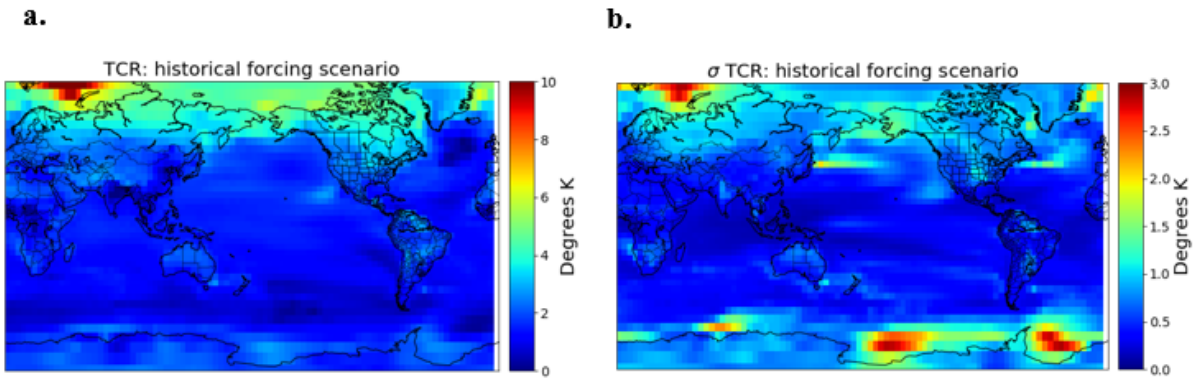

Figure 3.3: Ensemble average of the TCR estimated from historical forcing scenario (panel a) and the ensemble standard deviation of the TCR for historical forcing scenario (panel b).

Figure 3.4 shows the spatial variation of ΔT with respect to different variables within the Earth's energy budget. The ocean heat content (OHC) fraction below 100 meters, fOHCbelow100, is defined as the ratio of the change in OHC below 100 meters to the total change in OHC. High (low) groups of fOHCbelow100, λ , and total accumulated energy are constructed using the average of the 10 highest (lowest) ensemble members for each variable. For example, the ten members that have the highest fOHCbelow100 (change relative to 1850) are averaged to comprise the fOHCbelow100 high group.

The OHC fraction has an inverse relationship with TCR, the smaller OHC fraction members display larger ΔT at most latitudes (Fig. 3.4a, this relationship is discussed in detail in section 4). The maximum ΔT occurs at 70°N, with 3.0 K and 2.25 K warming in the low and high groups

respectively (Fig.3.4a). A secondary maximum is located at the South Pole and there is a local minimum in the Southern Ocean region. The zonal pattern of the difference between the maximum and minimum fOHCbelow100 members is a reflection of Earth's heat capacity. The largest differences between the maximum fOHCbelow100 group and the minimum fOHCbelow100 group occur in the midlatitude regions of both hemispheres. A near zero difference in ΔT occurs in the Southern Ocean region and in the tropics (Fig. 3.4b).

Following equation 1.2, the total accumulated energy in the large ensemble is equal to the sum of forced energy and the energy re-radiated to space. The total accumulated energy ($\int \Delta$Rdt) of the system is calculated by integrating the change in TOA flux imbalance over the entire globe for the length of the run. Further discussion of how each term is calculated and how Earth's energy balance controls the variability in the TCR is in section 4.

ΔT for the high and low λ groups and the fOHCbelow100 groups have similar warming patterns, with the most warming occurring in the polar regions. The high λ group peaks at approximately 70° N with a ΔT of about 3 K (Fig. 3.4c). Both the high and low λ groups are relatively constant in ΔT from the subtropical to the equatorial region (Fig. 3.4c). Figure 3.4d reflects this with a difference in ΔT of 0 K in the equatorial region, a 0.6 K difference at the South Pole, and a 1 K difference in the North Pole. Thus, λ can cause a large variability in ΔT focused in the polar regions. This is an artifact of the ice-albedo feedback, one component of the global feedback parameter.

The spatial pattern of the maximum and minimum accumulated energy groups is also investigated. The difference plot for total accumulated energy shows the largest difference between the groups at about 70°S latitude, but the magnitude of the difference is only about 0.23 K (Fig. 3.4e). In the northern hemisphere, the low accumulated energy group has a larger change in ΔT. Since the differences in the two groups are very small in magnitude, it is concluded that the total accumulated energy does not explain the variability observed in ΔT (Fig. 3.4f).

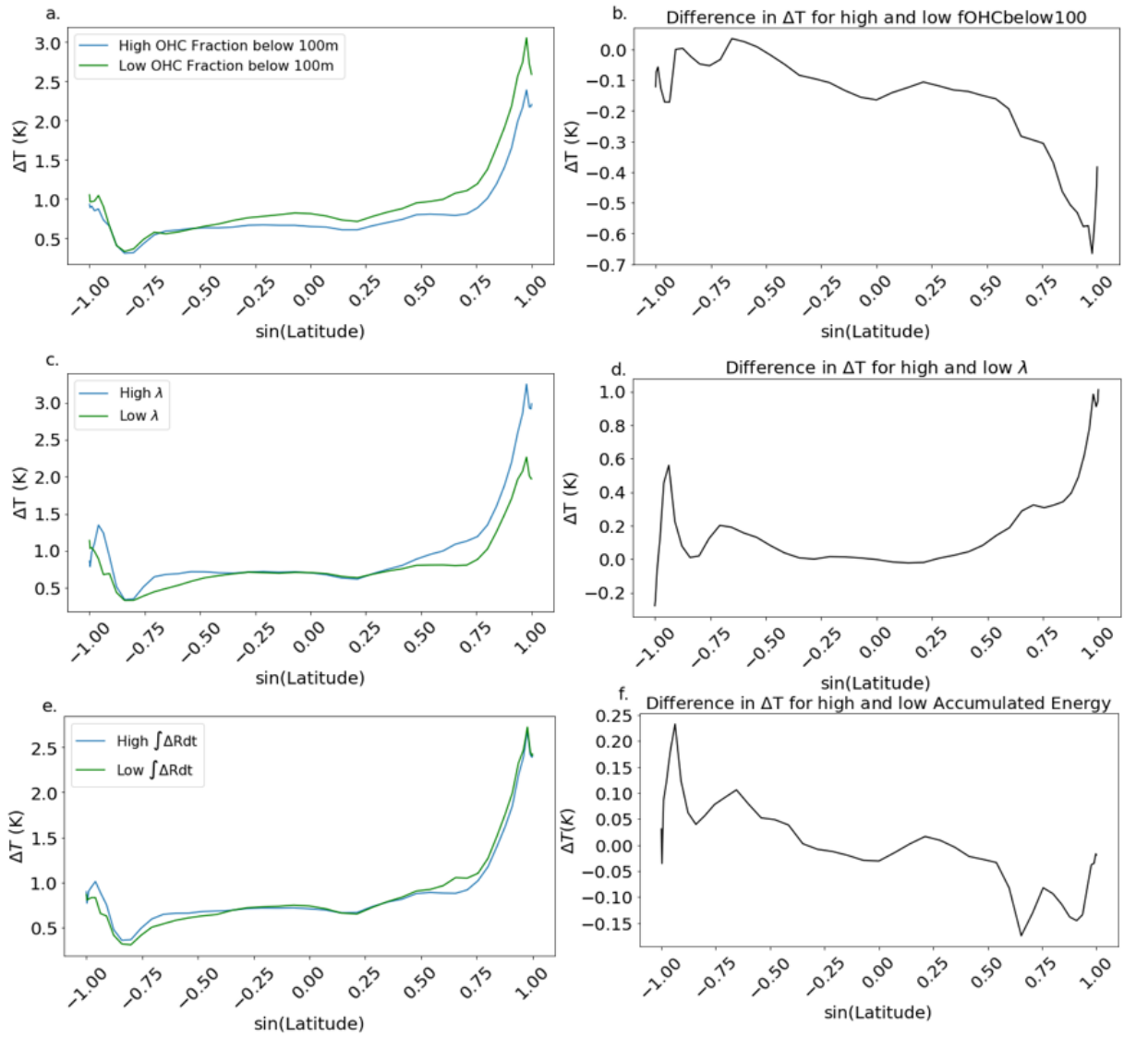

Figure 3.4: High (blue) and low (green) groups composed of the average of the 10 highest and 10 lowest TCR ensemble members respectively. The change in temperature (relative to 1859-1882 average) is plotted vs. the sine of latitude for both groups for fOHCbelow100 (panel a), λ (panel c), and total accumulated energy (panel e). The difference of the maximum and minimum TCR group is plotted against latitude in the left column (panel b, d, and f). A larger ΔT is representative of a larger spread in the data between the maximum and minimum groups for each variable.

The temperature change relative to the 1859-1882 base period is broken into the extratropical Northern Hemisphere (NH, 30°- 90°N), extratropical Southern Hemisphere (SH, 30°- 90°S), and tropical regions (TR, 30°S - 30°N). Using linear regression, ΔT for each region is correlated with

the TCR. In figure 3.5, it is found that a majority of the variability in the TCR can be explained by ΔT_{NH}. The R^2 value for the TCR for ΔT_{NH} is 0.82. The correlation coefficient for the TCR vs. ΔT_{TR} is 0.65. So, while a majority of the variability in the TCR can be explained by temperature changes in the NH, a portion of the variability in the TCR can also be explained by the linear relationship between TCR and ΔT_{TR}. The linear relationship between TCR and ΔT_{SH} has a low R^2 (0.32), and therefore does not contribute to explaining variability in the TCR.

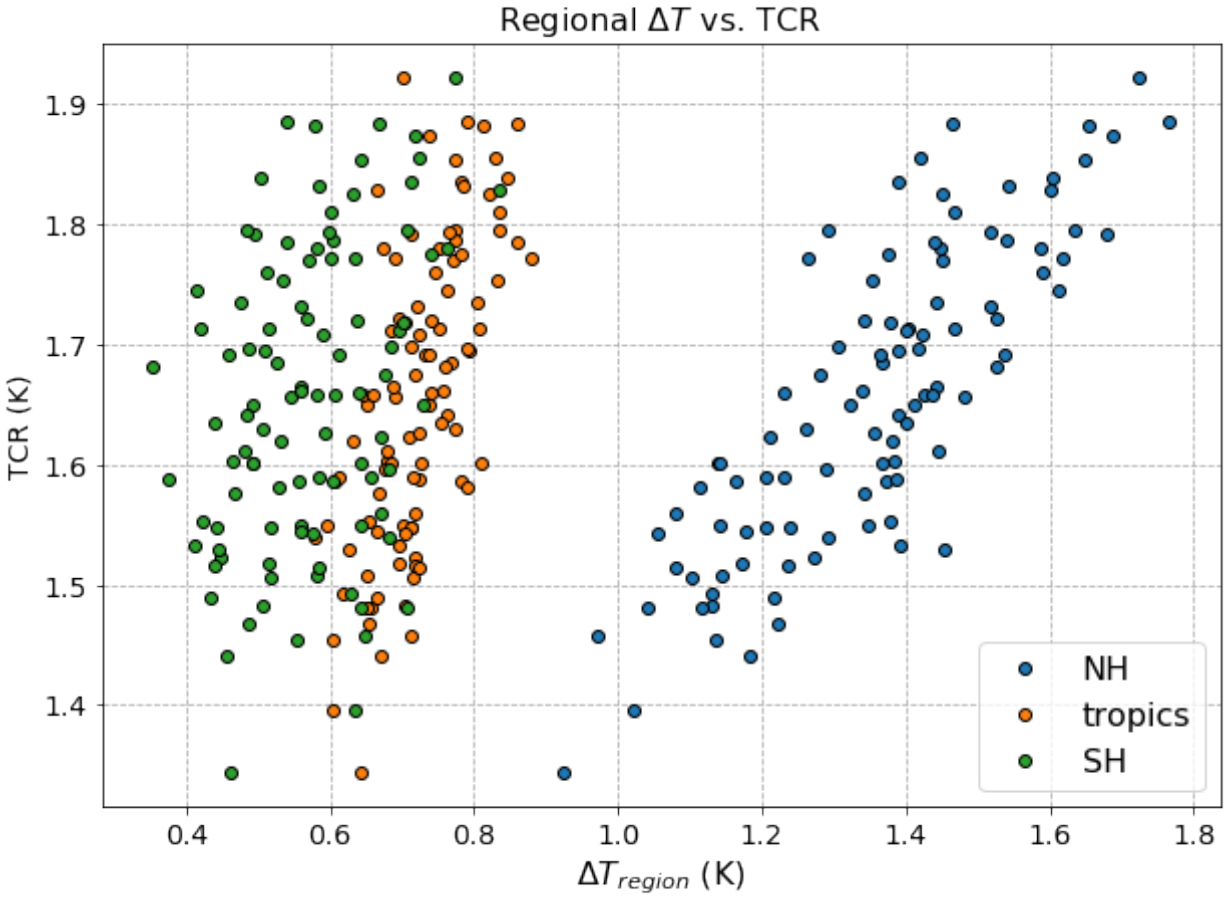

Figure 3.5: TCR vs. ΔT_{region} for the extratropical Northern Hemisphere (NH, blue), extratropical Southern Hemisphere (SH, green), and the tropical region (TR, orange). Linear regression is done and the correlation coefficients for linear regression with the TCR are found to be: $R^2_{NH} = 0.82$, $R^2_{TR} = 0.65$, and $R^2_{SH} = 0.32$.

3.3 Iteration of regions

To identify more precise regions that best explain the variability of the TCR, the average ΔT is calculated for each of 144 grid boxes using a 30° longitude x 15° latitude grid, for all 100 ensemble members. Our goal is to find a few combinations of 4 regional ΔT values that best explain the variability in the TCR. We use the following model:

$$TCR = aW + bX + cY + dZ \tag{3.1}$$

where W, X, Y, and Z represent the change in temperature for each individual ensemble member of four different grid points. We do this ordinary least squares (OLS) regression with all possible combinations of grid boxes, totaling about 17 million different possible models for the TCR. The groups of four regions that that yield the top four R^2 values are discussed below.

Figure 3.6a maps the 4 regions that yield the highest R^2 value with the TCR, in red (Fig. 3.6a). Together these four regions explain 79% of the variability in the TCR. The regions are located over the Arctic ocean (North of Alaska), the western coast of Europe, the southern tip of India, and in the Southern Ocean (Fig. 3.6a).

An ordinary least squares (OLS) regression using the set of 4 regions in figure 3.6b yields a R^2 value that is only 0.01% less than the highest R^2 (see figure 3.6a). Interestingly, the regions that comprise this nearly identical fit (in terms of R^2 value) are not completely different with those identified in figure 3.6a. The regions that comprise this model for the TCR are, the Arctic Ocean (north of western Russia), a region in Russia north of the Black and Caspian Seas, the maritime continent, and in the North Pacific Ocean (Fig.3.6b).

The regions used to yield the 3rd and 4th highest R^2 values are shown in figure 3.6c and figure 3.6d and explain 78.7% and 78.6% of the variability in the TCR, respectively. The regions used in these models are adjacent to those used for the second or first highest R^2 values (Fig. 3.6b-d).

All of the top 4 models of the TCR include similar regions (Fig. 3.6). The fits all include a region in the Arctic Ocean, a region in the Northern Hemisphere mid-latitudes near Russia or

Europe, a region in the North Pacific Ocean, and a region over the Maritime continent (Fig. 3.6b-d). Using ΔT in the four regions shaded in figure 3.6(a-d), OLS regression is able to explain about 79% of the variability in the TCR.

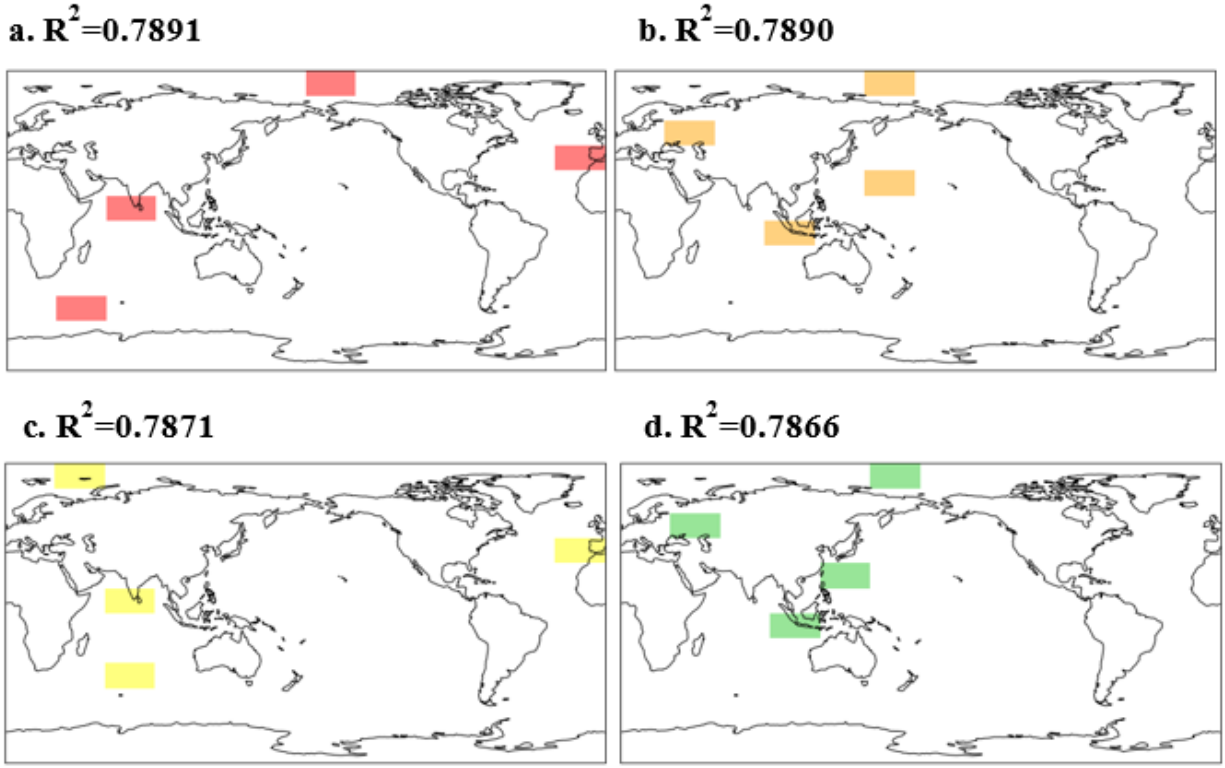

Figure 3.6: The average ΔT is calculated for each grid point within a 30° lon. x 15° lat. grid. ΔT for each grid box is then used as a variable in OLS regression, each grid box thus contains ΔT for all 100 ensemble members. The 4 grid boxes in which the ΔT yields the best fit for the TCR is shown in panel a. The R^2 value for the OLS fit using ΔT in the regions highlighted on the map is listed above each image.

Figure 3.7 shows the regions that yield the top 100 highest R^2 values in OLS regressions. The darker the shade of blue over a region means that the region appears in more than one of the top regressions. In general, the temperature change in regions throughout the Northern Hemisphere extra-tropics (specifically the Arctic and mid-latitudes) and the tropics consistently yield the highest R^2 values when used in OLS fit (Fig. 3.7). These regions agree with those found in in figure

3.5, the extra-tropical Northern Hemisphere and the tropics primarily control the variability in the TCR. Brown et al. (2016) also found that tropical Pacific and high latitude oceanic regions have a high correlation with the change in surface temperature. The variability in these regions can be attributed to oceanic currents and sea ice variability [Brown et al., 2016]. Future analysis using this method could be performed using regions over the ocean basins or larger grid boxes.

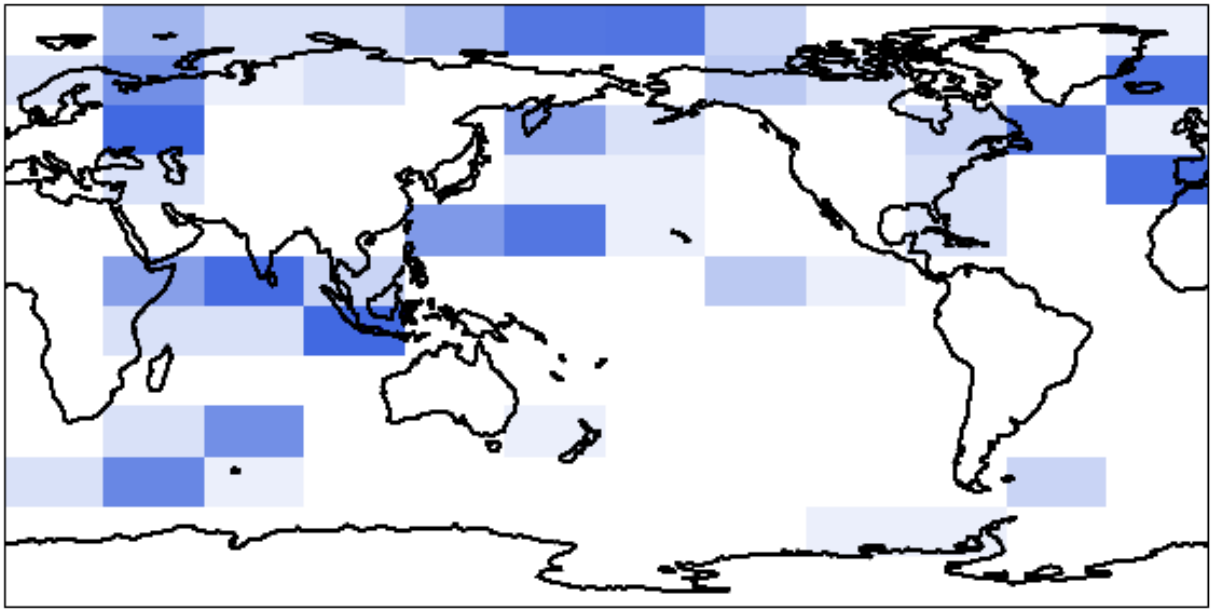

Figure 3.7: See caption for figure 3.6. This figure plots the regions that yield the top 100 highest R^2values, ranging from 0.76- 0.789. The darker shaded regions denote that a region is used in more than one OLS model (Eq. 3.1).

3.4 Conclusion

Using a 100-member large ensemble from the MPI Earth System Model (MPI-ESM1.1), we quantify the impact of internal variability on estimates of the TCR from the 20^{th} century historical record. We find that TCR estimates from individual members of the ensemble vary from 1.34 to 1.92 K, meaning that TCR estimated from any ensemble member may be strongly influenced by internal variability.

The spatial pattern of the TCR is investigated and it is found that a majority of the variability in the TCR occurs in the high latitude regions. Furthermore, when correlated with regional ΔT, we find that the temperature change in the extratropical Northern Hemisphere explains a majority of the variability in the TCR (82$\%$). Using a set of ΔT in four random regions in a 30° longitude x 15° latitude grid yields an OLS fit that can explain 79$\%$ of the variability in the TCR. The variability in the TCR can be explained by the temperature change in the Arctic Ocean (north of western Russia), a region in Russia north of the Black and Caspian Seas, the maritime continent, and in the North Pacific Ocean (Fig. 3.6b).

4. CONTROLLING FACTORS FOR OBSERVED VARIABILITY IN THE TCR

Earth's energy budget (Eq. 2.1) provides a physical framework from which the internal variability observed in the TCR can be explained. This section investigates the relationship of energy budget terms with the TCR. The fraction of ocean heat content (OHC), ocean heat content below 100 meters (fOHCbelow100), total OHC, OHC below 100, total accumulated energy, and λ are all correlated with the historical estimates of the TCR to determine what controls the internal variability observed.

4.1 Distribution of energy within the large ensemble

The total accumulated energy ($\int \Delta$R dt) of the system is calculated by integrating the change in TOA flux imbalance (shortwave plus longwave radiation) over the entire globe for the length of the run. The change in flux imbalance is calculated relative to the average of the TOA imbalance from a 2000-year control run. The energy leakage of the MPI-ESM2M1.1 model is 0.44Wm^{-2}, which is accounted for by subtracting the energy leakage from the total accumulated energy [Hedemann et al., 2017]. The energy re-radiated to space in the large ensemble ($\int \lambda\ \Delta$Tdt) is calculated by integrating the product of λ (described in section 2) and ΔT (relative to control run average) over the globe for the entire length of the run. The total forcing on the system is equal to the difference of the total accumulated energy and the energy re-radiated to space (Eq. 2.1).

Figure 4.1 depicts the relationship between total forcing in the system, energy re-radiated to space, and OHC. The total forcing of the system between 1850 to 2005 is about 107x10^{22} J, and the uncertainty range varies from 93-119x10^{22} J (5-95%). The extra energy radiated out to space for 1850-2005 is about 63.2x10^{22} $\pm$ 48- 77 x 10^{22} J, and the total accumulated energy (about equal to OHC) is 42.7x10^{22} $\pm$ 38-48x10^{22} J (Fig. 4.1).

Huber and Knutti (2012) use probabilistic estimates from an intermediate complexity climate model ensemble to quantify the magnitude of these components within the energy budget. The total accumulated net forcing (1850-2010) energy varies from 95-197 x 10^{22} J (5-95% range) and

the 5-95% range of net ocean heat uptake is 125 to 275 x 10^{22} J [Huber and Knutti, 2012]. Murphy et al. (2009) quantifies the energy budget from 1950 to 2010 using observations and find that energy radiated to space ranges from 200 to 550 x 10^{22} J and about 200x10^{22} J is stored in the ocean. The overall distribution of energy agrees with past energy balance studies in the fact that about of forcing energy in the large ensemble is radiated back to space and about one third of it is stored as accumulated energy (Fig. 4.1) [Murphy et al., 2009, Huber and Knutti, 2012]. About 90-93% of Earth's energy imbalance ($\int \Delta R\, dt$) is sequestered into the ocean as heat uptake [Von Schuckmann et al., 2016;Allan, 2017]. In the large ensemble 82-95% (5-95%) of the total accumulated energy is stored in the ocean, with an ensemble average of 89% . Thus, large ensemble energy balance agrees with previous estimates of the Earth's energy balance, despite different data sets and periods of analysis.

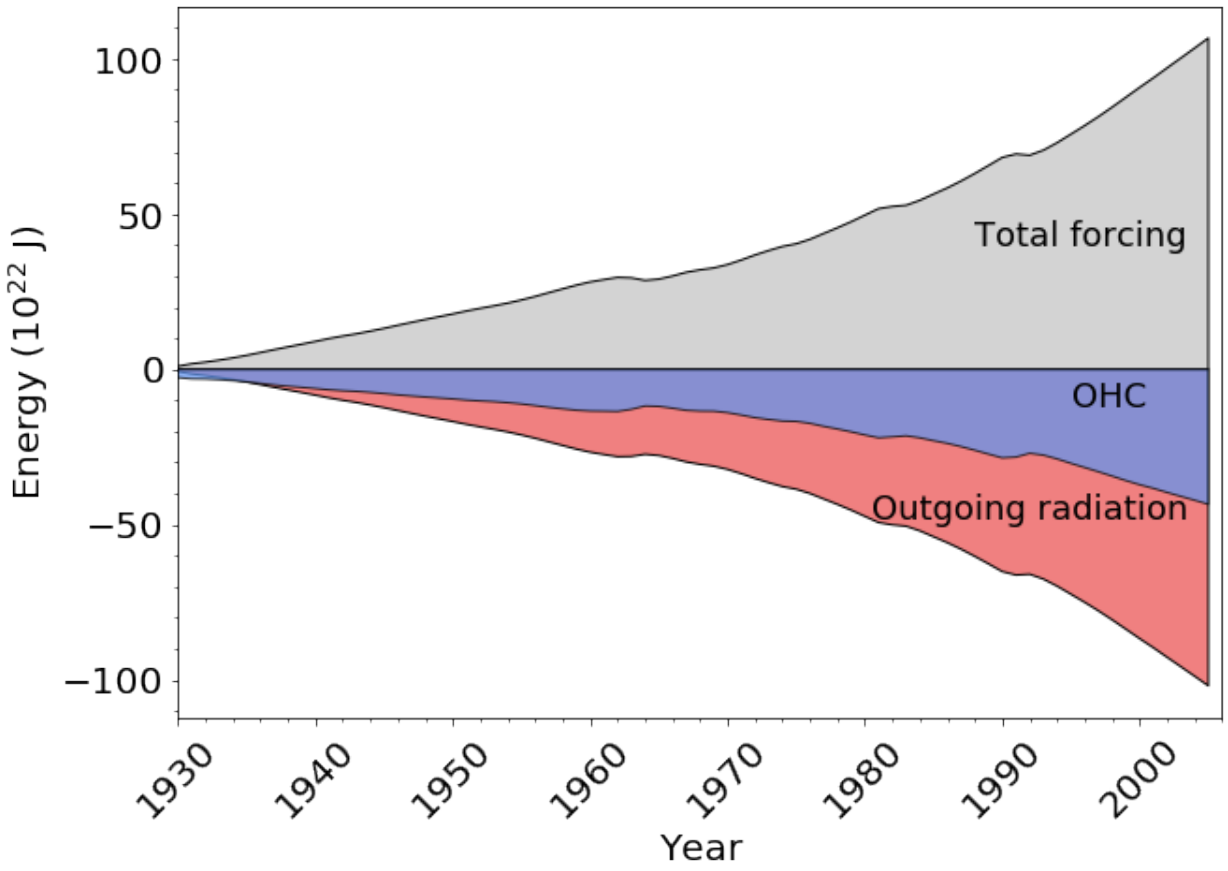

Figure 4.1: Adapted from Huber and Knutti (2012), figure 2d. The ensemble average outgoing radiation ($\int \lambda \Delta T dt$), the ocean heat content (assumed to be proportional to $\int \Delta R dt$), and total forced energy ($\int \Delta F dt$) are calculated from the large ensemble. The Δ terms are relative to the large ensemble pre-industrial control run.

4.2 Selecting the depth of OHC fraction

We analyze the OHC stored below a range of ocean depths from 17 to 5720 meters. The change in OHC over the length of the run for each depth is calculated by taking the difference between the 1995-2005 average and the 1859-1882 base period (same method as ΔT). The OHC fraction is defined as the ratio of the change in OHC below each depth to the total change in OHC.

Linear regression is done with the OHC fraction and the TCR at each depth for each member of the large ensemble. The results of the correlation vs. ocean depth are shown in figure 4.2. As expected, the correlation coefficient of OHC fraction with TCR decreases as depth increases

(Fig. 4.2). There is a relatively constant r value in the mixed layer, the r value declines below 1000 meters, and then the correlation coefficient approaches zero in the deep ocean. This result makes sense, the high-density water deep below the ocean surface has a negligible relationship with the transient warming at the surface. Following equation 1.1 and the conservation of energy, as more heat is sequestered into the ocean as OHC, less surface warming is realized. As the OHC below 100 meters (fOHCbelow100) increases, the TCR decreases; the r-value explains this inverse relationship with a value of -0.8 (Fig. 4.2). Due to the strong relationship with TCR the fOHCbelow100 is chosen as the ocean heat content variable for analysis, it also represents the climatological mixed layer depth.

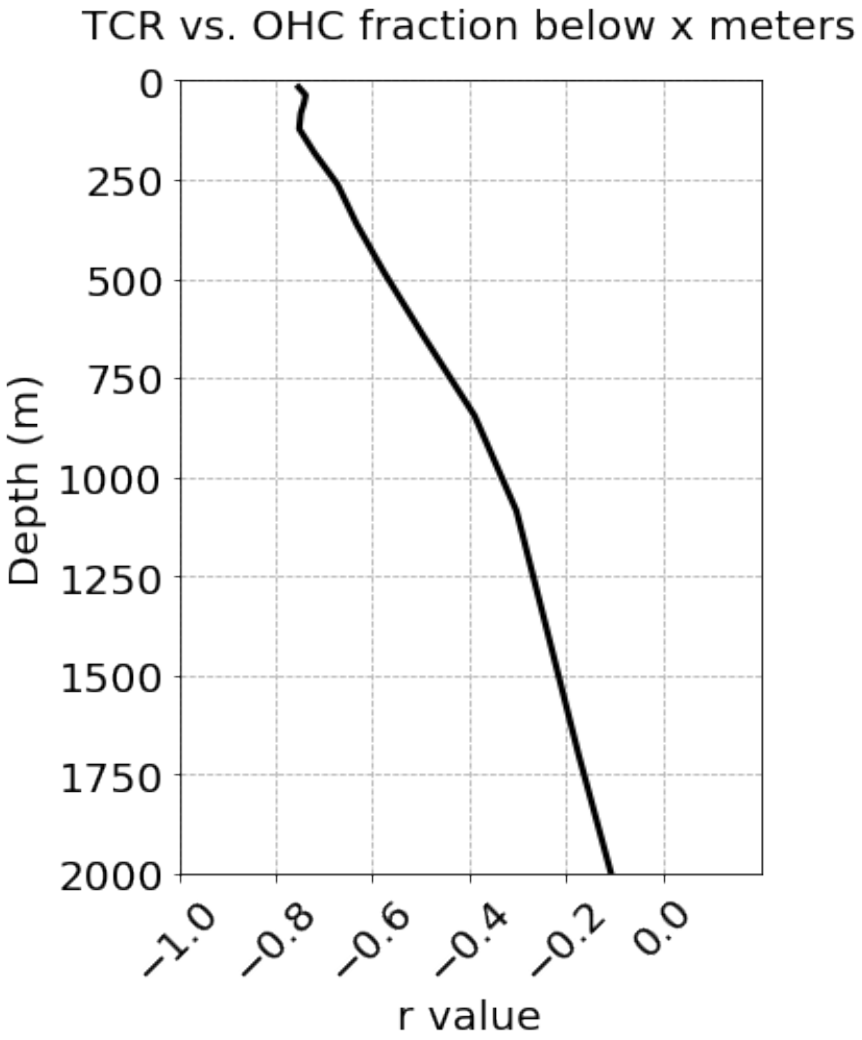

Figure 4.2: Plot of the correlation coefficient of TCR vs. OHC fraction below ocean depths from 17 meters to 2000 meters.

4.3 Distribution of energy balance terms

The magnitude of each term within the linearized energy budget is examined to characterize the distribution of energy within the large ensemble. Figure 4.3a shows that 81-87% of the total OHC is stored below 100 meters (5-95%) in the large ensemble. The variables that comprise this ratio, OHC below 100 meters (Fig. 4.3c) and the OHC total (Fig. 4.3b) have average values of 3.14 x10^{23} J and 3.8 x10^{23} J, respectively. The OHC total is about 90% of the total accumulated

energy (10^{23} J) of the system, and the mean is 4.3x10^{23} J (Fig. 4.3d). The energy reradiated out to space ($\int \lambda \Delta T dt$) has a mean of 6.4 x10^{22} J. The climate feedback parameter, λ , controls the rate at which energy is radiated out to space and ranges from -1.55 to -1 $Wm^{-2}K^{-1}$ in the large ensemble (Fig. 4.3f).

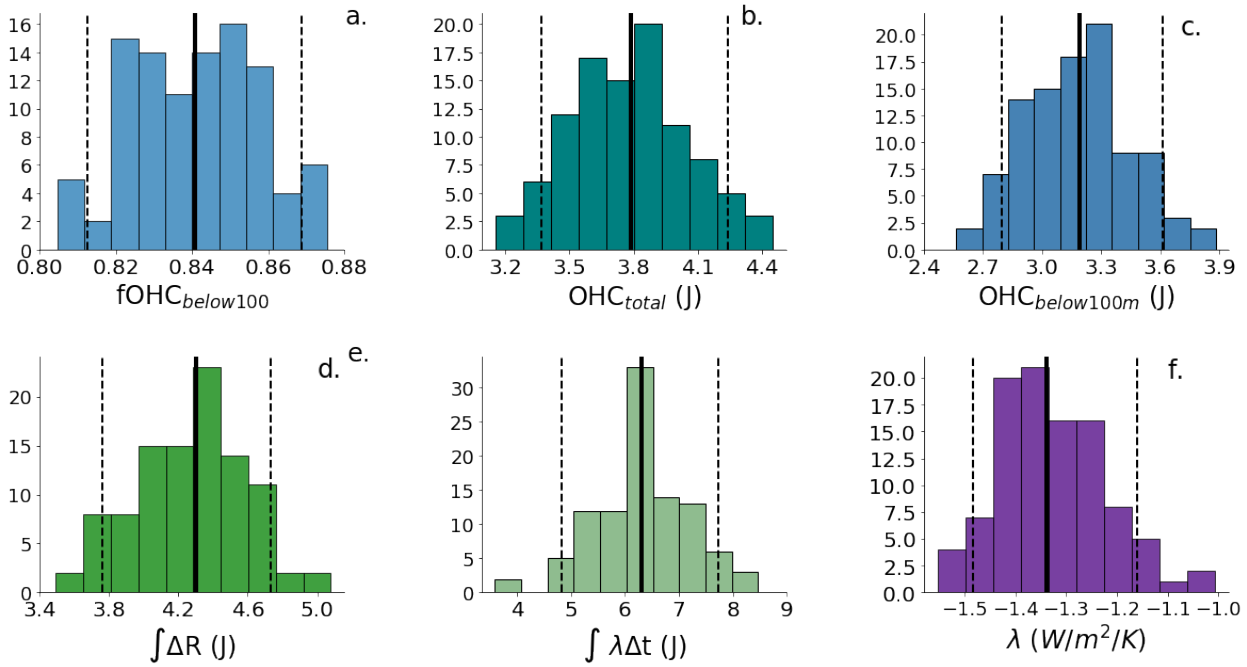

Figure 4.3: Distributions of OHC fraction below 100 m (a), total ocean heat content (10^{23} J) (b), ocean heat content below 100 meters (10^{23} J) (c), accumulated energy (10^{23} J), (d) energy radiated to space (10^{23} J) (e) λ ($Wm^{-2}K^{-1}$). The solid line represents the variable mean, and the dashed lines represent the 5-95% range.

4.4 Correlation of energy budget variables with the TCR

Previous studies have stated that ocean depths below the mixed layer are important in controlling the climate sensitivity [Armour et al., 2012; Held et al., 2010; Raper et al., 2002]. Armour et al. (2012) states that "The uptake of heat by the deep ocean strongly influences transient warming by acting as a sink of energy at the surface". In the large ensemble, we find a strong linear relationship between TCR and the fraction of energy stored in the ocean below 100 m (R^2= 0.57)

(Fig. 4.3a). The negative correlation shows that as model members store more heat in the deep ocean they tend to have a lower TCR, which follows the concept of conservation of energy since each ensemble member has the same amount of radiative forcing.

The total OHC (J) and the OHC below 100 meters (J) have much weaker linear correlations with the TCR (Fig. 4.4b and 4.4c). The correlation coefficients are less than 0.05, and p-values show that the correlations are not statistically significant at the 95% level (Table 4.1).

The fraction of the OHC below 100 meters comprises 81-87% (5-95th percentile) of the total OHC, yet it can explain a significant portion of the variability in the TCR that the total OHC and OHC below 100 meters in Joules cannot. It is hypothesized that both the OHC and OHC below 100m have larger variability than the fOHCbelow100, which causes a weak linear correlation between these variables and the TCR. OHC total and OHC below 100 meters have variance of about 0.08 J^2 and the variance of fOHCbelow100 is 2 orders of magnitude smaller (Table 4.1). The coefficient of variation is a unitless measure of dispersion that divides the standard deviation by the sample mean. The CV are 8% , 7% , and 2% for OHC below 100m, OHC total, and fOHCbelow100, respectively (Table 4.1). This means that OHC and OHC below 100m have several times the variability of fOHCbelow100, and this variability tends to obscure any correlation that might exist with TCR.

The climate feedback parameter (λ) for each ensemble member is calculated using the linearized energy balance equation (Eq. 1.1) [Dessler et al., 2018]. Model members with a given value of TCR display a broad range of λ , yet the linear relationship between the two explains 21% of the variability in the TCR (Fig. 4.4f). The very small p-value is indication that the linear correlation between TCR and λ is statistically significant at the 95% level (Table 4.1). The total accumulated energy ($\int \Delta R$ in Eq. 2.1) explains about 1% of the variability in the TCR and the linear correlation is not statistically significant (Fig. 4.4f, Table 4.1). The energy re-radiated to space ($\int \lambda \Delta T$ in Eq. 2.1has a slight negative correlation with the TCR, explains about 7% of the variance, and the relationship with the TCR is statistically significant (Fig. 4.4e, Table 4.1).

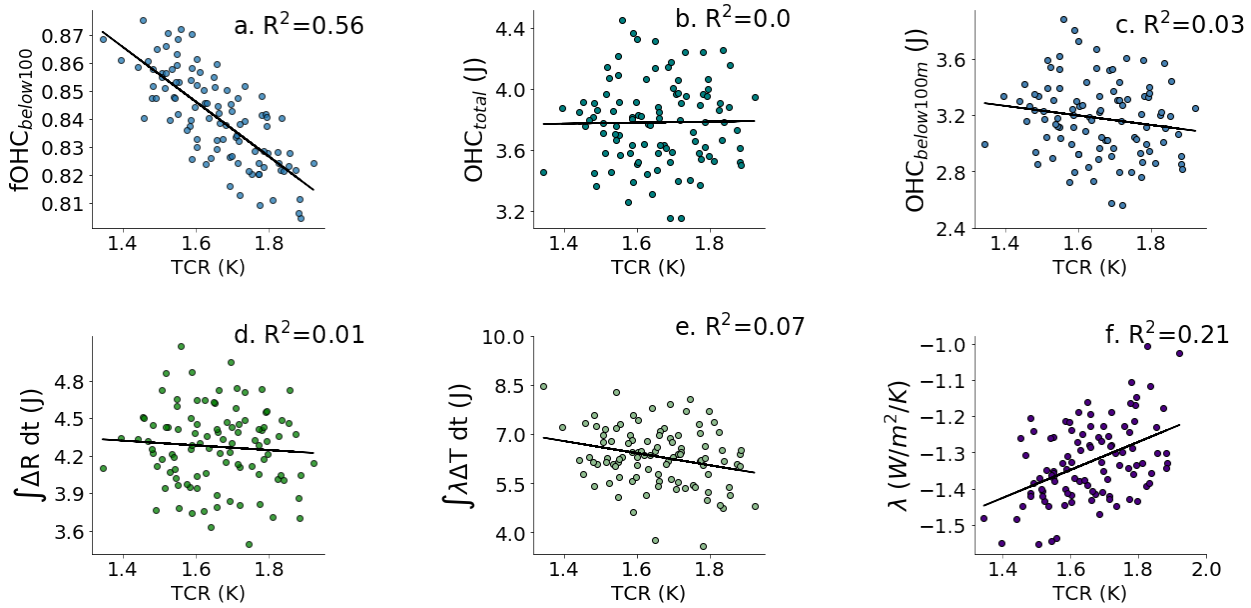

Figure 4.4: Scatter plot for 100 ensemble members of TCR vs. (a) OHC fraction below 100 m, (b) total ocean heat content (10^{23} J), (c) ocean heat content below 100 meters (10^{23} J), (d) accumulated energy (10^{23} J), (e) ($Wm^{-2}K^{-1}$). Linear regression is over-laid in the thin black lines and the legends display the correlation coefficient values for each regression.

Variable Correlated with TCR	**R^2**	**P-value**	**Variance (Units squared)**	**CV (Unitless)**
OHC Fraction below 100m	0.564	2.21×10^{-19}	2.7×10^{-4}	0.020
OHC Total (J)	2E-4	0.88	0.075	0.072
OHC below 100m (J)	0.026	0.11	0.072	0.084
Accumulated Energy (J)	0.006	0.45	0.099	0.074
Energy Radiated (J)	0.067	0.009	0.0081	0.14
λ (Wm^{-2}/K)	0.21	2.24×10^{-6}	0.0115	-0.08

Table 4.1: R^2 and p-value of all the linear correlations in Fig. 4.4 of energy balance variables with the TCR. Variance and covariance of variables are listed in the last two columns.

4.5 Multivariate linear regression

We now explore whether using two variables yields a better fit for the TCR. After some trial and error, it is found that the total energy re-radiated and the OHC fraction below 100 m control a large portion of variability in the TCR (Fig.4.5). The OHC fraction is clearly the dominant explanation for the observed variability in the TCR (Fig.4.4a, Fig.4.5). The TCR decreases as the OHC fraction below 100 meters increases, and the relationship of TCR with energy re-radiated to space is less distinct (Fig. 4.5).

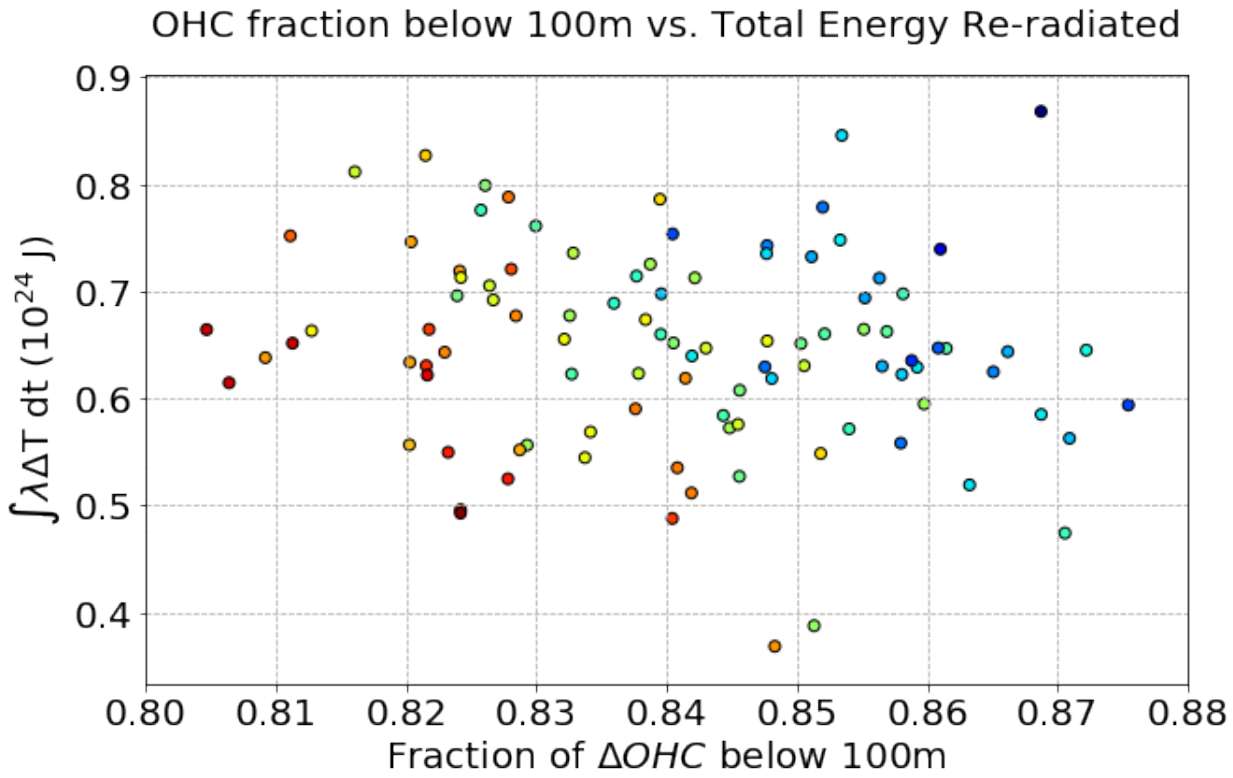

Figure 4.5: Scatter plot for 100 ensemble members of TCR vs. (a) OHC fraction below 100 m, (b) total ocean heat content (10^{23} J), (c) ocean heat content below 100 meters (10^{23} J), (d) accumulated energy (10^{23} J), (e) λ ($Wm^{-2}K^{-1}$). Linear regression is over-laid in the thin black lines, and the legends display the correlation coefficient values for each regression.

The results from figure 4.5 provide a qualitative idea that we can get a more accurate estimate

of TCR using two parameters. OLS regression is used to quantify the prediction of the TCR with two parameters (Eq. 4.1).

$$TCR = a * X + b * Y \tag{4.1}$$

Results of several ordinary least squares (OLS) regression fits with two parameters (X and Y) to model the TCR are shown in Table 4.2. It is found that total energy re-radiated to space and fOHCbelow100 control a maximum of the variability in the TCR, the R^2 is 0.71 (Fig. 4.6). The model that includes the total energy accumulated and the OHC below 100 meters explains about 69% of the variability observed in the TCR (Table 4.2).

The root mean squared error (RMSE) measures how much the predicted values from the OLS fit of the TCR deviate from the observed values. The RMSE for the model using energy radiated to space and OHC fraction below 100m is 0.067 K and the RMSE for total accumulated energy and fOHCbelow100 is 0.069 degrees K (Table 4.2).

X	**Y**	**R^2**	**P- val**	**RMSE (degrees K)**
Total Accumulated Energy	OHC fraction below 100m	0.69	1.08 x10^{-26}	0.069
Total Accumulated Energy	OHC below 100 m	0.035	0.06	0.123
Total Accumulated Energy	OHC total	0.024	0.12	0.123
Total radiated energy	OHC fraction below 100m	0.71	6.5 x10^{-28}	0.067
λ	OHC fraction below 100 m	0.65	8.87 x10^{-24}	0.074

Table 4.2: The results of OLS fits with the TCR and two parameters (variable 1 and variable 2 columns in chart). The correlation coefficient, p-value, and RMSE of each OLS fit prediction of the TCR with the observed TCR values.

The best fit for TCR is:

$$TCR = a * fOHCbelow100 + b * \int \lambda \Delta T dt \tag{4.2}$$

The coefficients a and b are -6.1 $\pm$ .0.42 K and 5.2x$10^{-25}\pm$ 7.5x10^{-26} K $(Wm^{-2})^{-1}$ (uncertainty bounds are one standard error).

Since energy trapped by forcing ($\int \Delta$F, Eq. 1.2 is equivalent for every ensemble member, variability in the total energy accumulated ($\int \Delta$Rdt) is about equal to variability in the energy radiated to space ($\int \lambda \Delta$T dt). This explains the similarity between the OLS models using total accumulated energy and total energy radiated to space (Table 4.2).

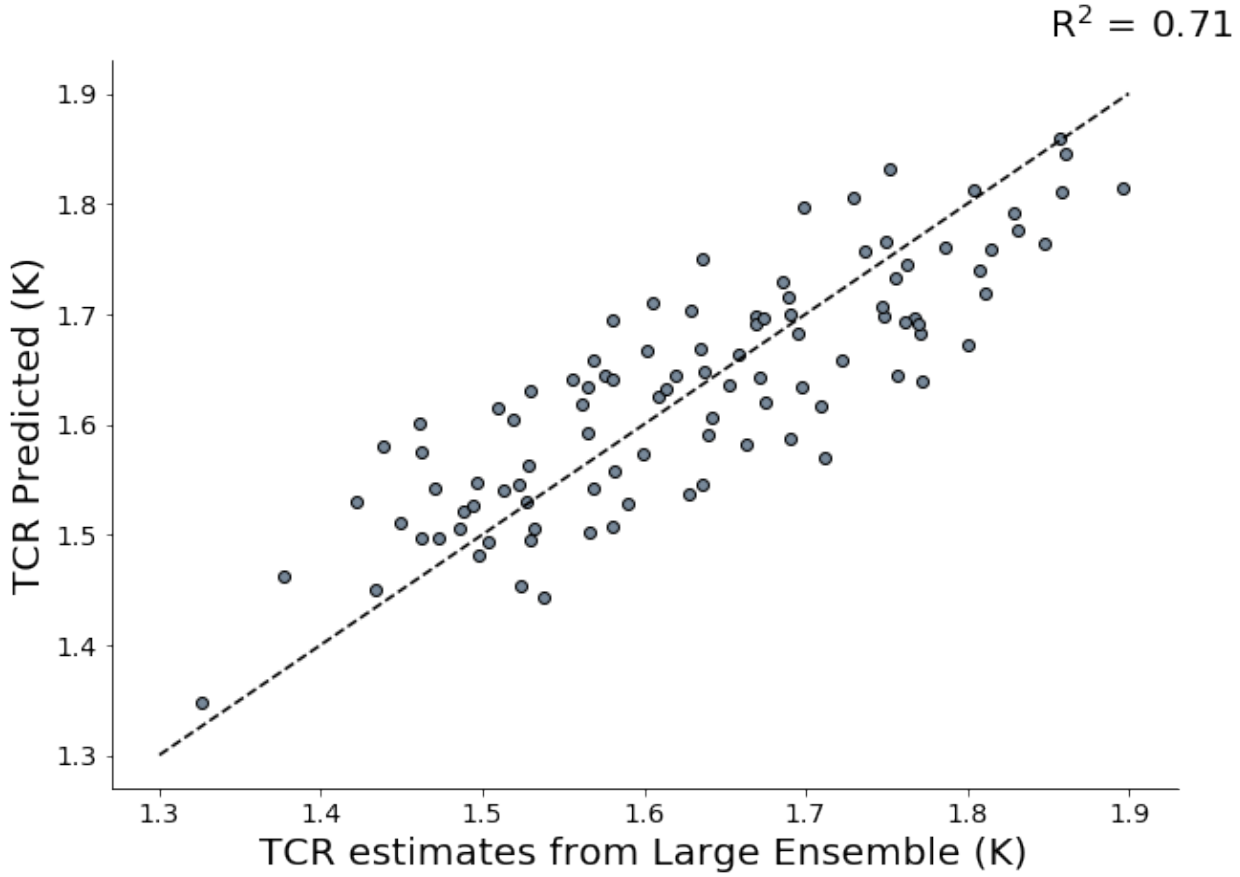

Figure 4.6: TCR predicted by the ordinary least squares fit (Eq. 4.2) with OHC below 100 meters and total radiated energy (y-axis) and TCR estimates from the large ensemble (x-axis). The R^2 value of this fit is 0.71.

4.6 Conclusion

The Earth's linearized energy budget equation provides a framework with which the observed internal variability can be analyzed (Eq. 1.1). We find that TCR correlates most strongly with the

fraction of OHC that is stored below 100 meters (R^2 of the correlation is 0.57). Thus, the model members that sequester more energy in the deep ocean show less warming and therefore have a lower TCR.

Using two variables fraction of energy stored below 100 m and energy radiated back to space provides an even better fit (explaining 71% of the variability).This is consistent with energy conservation: the more energy that is stored in the deep ocean or radiated back to space, the less surface warming the Earth experiences.

5. TIME DEPENDENCE OF THE TCR

Previous studies have examined the time dependence of the TCR, by comparing ΔT from the first to the second doubling of CO_2. It is assumed that F_{2xCO2} is proportional to the logarithm of the forcing and therefore increases at a constant rate for each successive doubling of CO_2 [Gregory and Forster, 2008]. Several previous studies have found that ΔT from the first doubling of CO_2 is smaller than the ΔT in response to the second doubling of CO_2 (hereafter referred to as TCR2) [Gregory and Forster, 2008; Gregory et al., 2015]. Thus, the rate of temperature change is not constant with respect to forcing in the 1% per year scenario [Gregory and Forster, 2008; Gregory et al., 2015].

5.1 ΔT in the 1% per year forcing scenario

Figure 5.1 shows the temperature increase in the 1% / yr. forcing scenario (starting in 1850), for all 68 ensemble members. Relative to constant dT/dt (Fig. 5.1, dashed black line), the temperature increases at an accelerating rate throughout the run. The rate of ΔT increase deviates the most from the constant dT/dt at the beginning and end of the simulation (Fig. 5.1). The slope of ΔT at the beginning of the run is shallow compared to a constant dT/dt and the end of the run has an accelerated rate of warming compared to constant, the middle of the run has a constant rate of ΔT increase (dashed line). North and Kim (2017) state that linearly forced systems are likely to asymptote to a linear increase in temperature after a time lag. Figure 5.1 displays this behavior, the beginning portion of the run displays non-linear growth and then the ensemble average change in temperature asymptotes to a linear increase in temperature with time (North and Kim, 2017).

Related to the non-linear increase in temperature change with time, figure 5.1 shows that ΔT centered around the first doubling of CO_2 is smaller than the ΔT for the quadrupling of CO_2. The average ΔT is 1.8 K for the first doubling (year 1920), and 2.4 K for the second doubling of CO_2.

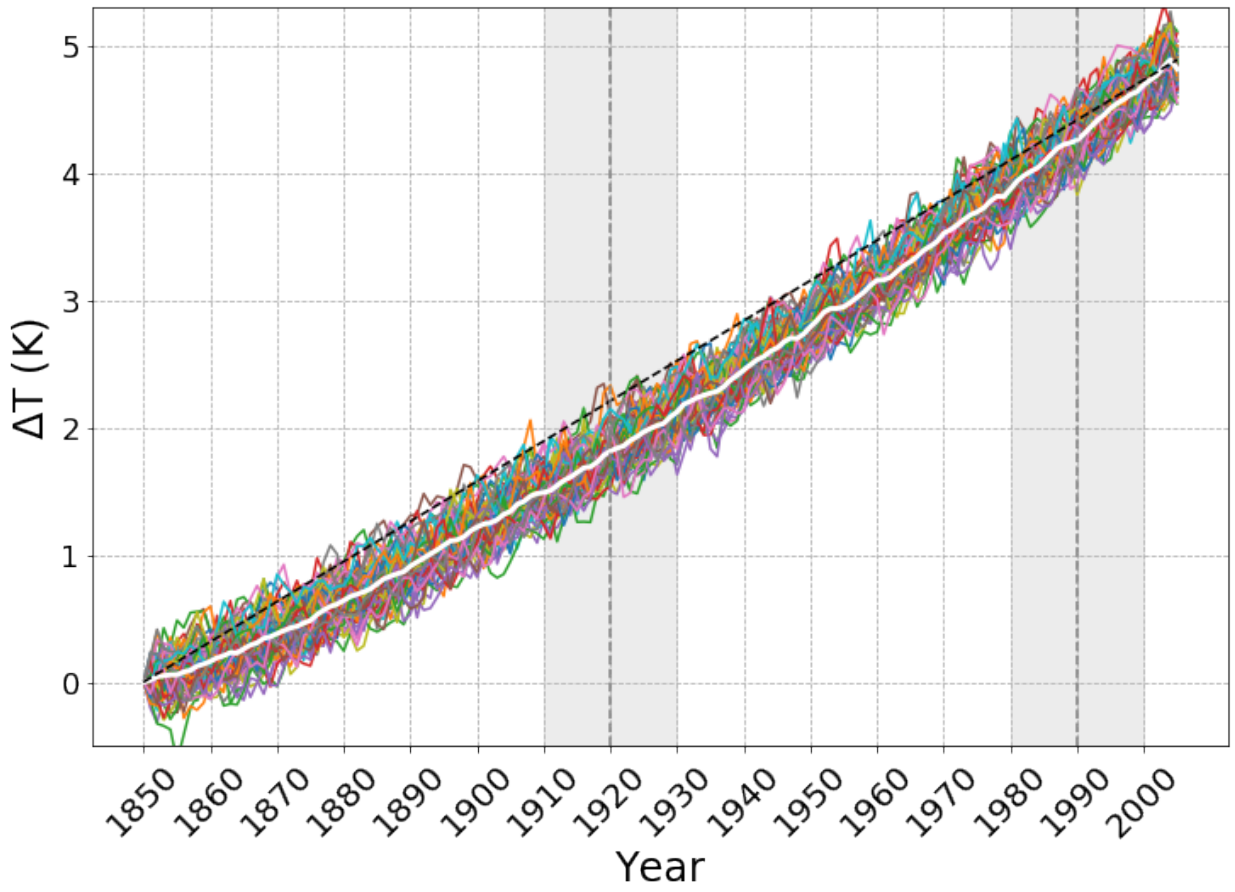

Figure 5.1: Adapted from Gregory et al. (2015), figure 4. Using data from the large ensemble 1% per year forcing scenario, annual ΔT relative to the average of year 1850. Colored lines represent individual ensemble members (68-member ensemble) and the white line represents the ensemble mean. The periods of CO_2 doubling (year 70) and CO_2 quadrupling (year 140) are shown (dashed grey lines). The periods centered around the time of 2xCO_2 and 4xCO_2 are lightly shaded. The black dashed line represents a constant change of temperature with time.

5.2 First vs. second doubling of CO_2

Equation 5.1 is used to quantify the differences in ΔT between the first and second doubling of CO_2 within the large ensemble. To avoid confusion with the estimates of the TCR made from the historical ensemble, we refer to the warming centered around year 70 in the 1% per year increase in CO_2 scenario as the TCR1 and the warming centered around year 140 as the TCR2. Following the method of Gregory et al. (2015), TCR1 and TCR2 are calculated using a 20-year mean centered around the time of CO_2 doubling and quadrupling (Eq. 5.1).

$$TCR1 = \overline{1910 - 1930} - \overline{1850}$$
$$TCR2 = \overline{1980 - 2000} - \overline{1850} \tag{5.1}$$

In the case that the TCR is constant in time, the ratio of TCR1 to TCR2 would be 0.5 [Gregory et al., 2015]. Figure 5.2 shows the distribution of TCR1/TCR2 for all 68 model members (forced by the 1% per year increase scenario). The ratio ranges from 0.38 to 0.47 and the mean is about 0.42. Gregory et al. (2015) obtained a range of 0.38 to 0.47 from the use of 16 CMIP5 AOGCMs. Our range of TCR1/TCR2 is therefore consistent with the ratio observed in CMIP5 models.

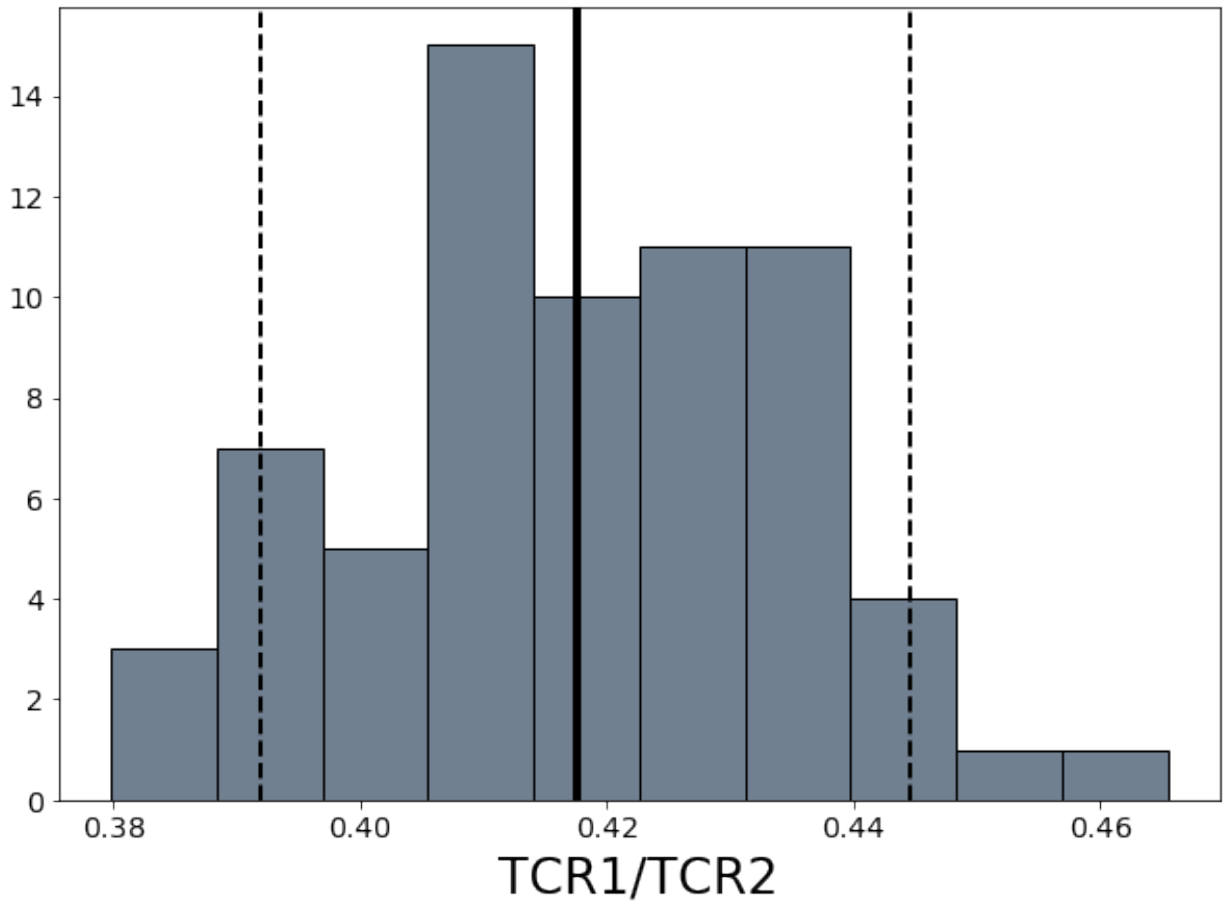

Figure 5.2: The ratio of TCR1 to the TCR2. The TCR1 is defined to the be the temperature change at the time of CO_2 doubling in a 1% per year increase scenario. TCR2 is calculated as the temperature change in response to the second doubling of CO_2.The solid black line represents the mean, and the dashed lines represent the 5-95 percentile.

It is hypothesized that model members that have a large magnitude of warming in the first 70 years of the simulation (relative to the ensemble mean) will have less warming in the second half of the simulation relative to average. Since TCR1 and TCR2 both contain the period from preindustrial to the first doubling of CO_2, a new metric is defined to ensure independence. $\Delta T_{2\text{-}4xCO2}$ represents the temperature change due to the second doubling of CO_2 relative to the first doubling of CO_2 (essentially TCR2-TCR1 in equation 5.1). On average the $\Delta T_{2\text{-}4xCO2}$ is 0.7 K larger than the TCR, the fact that $\Delta T_{2\text{-}4xCO2}$ is greater than TCR1 for all members of the large ensemble agrees with past studies [Gregory and Forster, 2008; Gregory et al., 2015; Grose et al., 2018].

There is a slight negative trend present in the plot of $\Delta T_{2\text{-}4xCO2}$ vs. the TCR1 (Fig. 5.3). The slope of $\Delta T_{2\text{-}4xCO2}$ and the TCR1 is -0.09, representing an inverse relationship between $\Delta T_{2\text{-}4xCO2}$ and the TCR1.The correlation coefficient between the two is -0.24, and the p-value is 0.05. Thus, there is sufficient evidence to support that there is a relationship between the magnitude of warming for the first doubling of CO_2 and the warming observed for the second doubling of CO_2. While the relationship is statistically significant, the correlation coefficient is evidence that the inverse relationship is not robust.

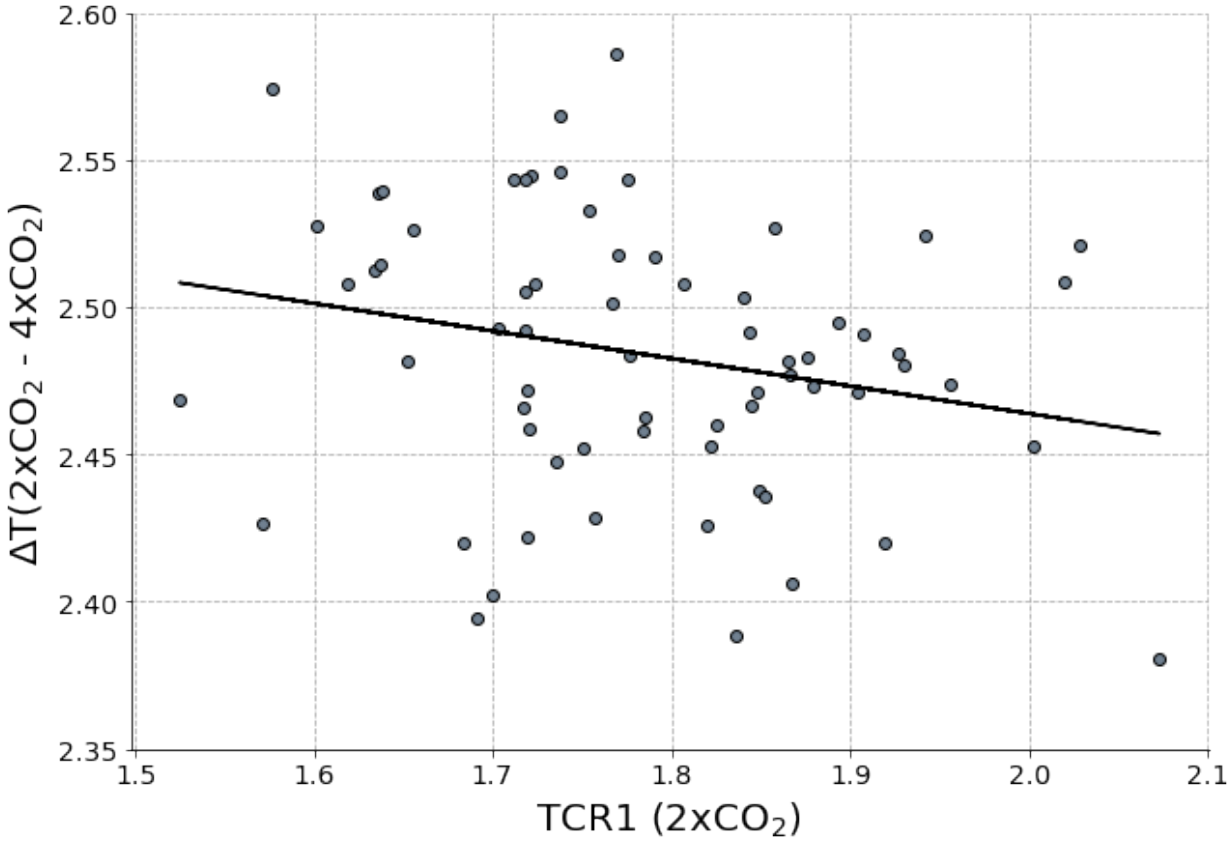

Figure 5.3: ΔT (relative to 1850) for the second doubling of CO_2 in the 1% per year CO_2 increase scenario vs. the TCR. The black line represents a linear regression of the two metrics.

5.3 Spatial pattern

5.3.1 Ratio of $\Delta T_{2\text{-}4xCO2}$ to the TCR

Figure 5.4 shows the spatial pattern of the ratio of TCR1 to $\Delta T_{2\text{-}4xCO2}$ and the standard deviation of the ensemble, specifically zonal means at intervals of 4 degrees latitude. The spread of the ratio in the ensemble displays a maximum at about 70°S latitude (sin(latitude) = 0.95, Fig.5.4a). The standard deviation of the ratio remains below 0.15 for all latitudes north of the maximum. There is a local maximum at the equator, with a standard deviation of 0.15 K. Thus, a majority of the variability in the ratio of TCR1 to $\Delta T_{2\text{-}4xCO2}$ is located at the intersection of the Southern Ocean and the Antarctic edge, and there is some contribution from the equatorial region which is likely due to El Nino Southern Oscillation or an increased mixed layer depth.

The average ratio of the ensemble is maximized in the mid-latitude regions, with average values of 0.8 at 70°S and 0.9 at about 80°N. (Fig. 5.4b). These latitudes represent the regions where there is the smallest difference between the TCR and $\Delta T_{2\text{-}4xCO2.}$ These regions thus have the least amount of delayed response in their warming. In the midlatitudes, the ratio is minimized at about 55°S in the Southern Ocean region with a value of about 0.6. In the tropical region (30°S to 30°N) the average ratio remains at about 0.7 (Fig. 5.4b). These regions display a larger delay in their temperature changes, which could be due to the fact that these regions have higher heat capacities, so they take longer to respond to a given forcing.

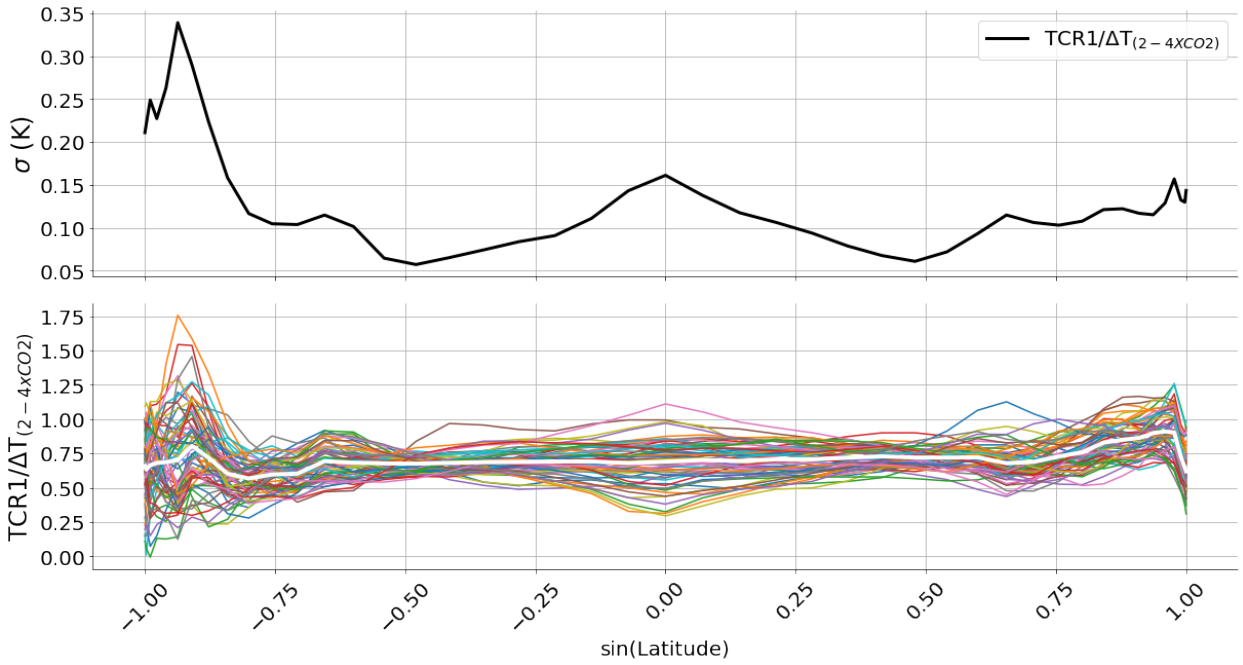

Figure 5.4: Panel a: Standard deviation of the 68-member ensemble for the ratio of $\Delta T_{2\text{-}4xCO2}$ and TCR vs. the sine of latitude. Panel b: Plot of ensemble ratio vs. the sine of latitude. The white line is the ensemble average.

5.3.2 Difference of $\Delta T_{2\text{-}4xCO2}$ and the TCR

All model members display more warming during the second doubling of CO_2 ($\Delta T_{2\text{-}4xCO2}$) than the first doubling (TCR1). The TCR1 is maximized at 80° N, and then it decreases to the North Pole (Fig. 5.5). The $\Delta T_{2\text{-}4xCO2}$ has a similar warming pattern but diverges at about 75° N, where it continues to increase in warming through the entire Arctic region. This causes a large difference between the two quantities in the North Pole region, $\Delta T_{2\text{-}4xCO2}$ is 2.7 K greater than the TCR. The zonal pattern of the difference between $\Delta T_{2\text{-}4xCO2}$ and the TCR1 is likely a consequence of the ice-albedo feedback. The difference decreases in the northern mid-latitude region and remains less than 1 K throughout the tropics and southern mid-latitude region. A minimum in the difference is located at about 65°S, both the TCR1 and $\Delta T_{2\text{-}4xCO2}$ have minimal difference of warming in this high heat capacity region. Another local maximum in the difference occurs at the South Pole and is about 1.25 K (Fig. 5.5).

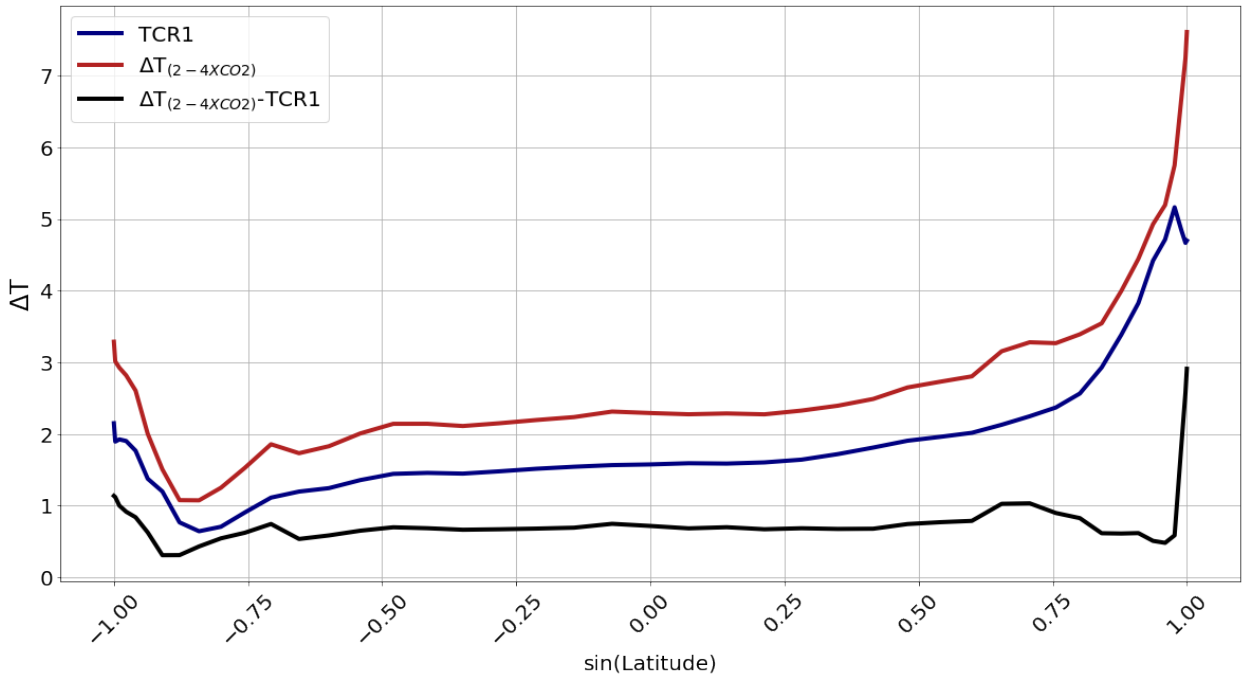

Figure 5.5: The zonal pattern of the TCR1 (blue), $\Delta T_{2\text{-}4xCO2}$ (red), and the difference of the two (black). All quantities represent the ensemble average.

The spatial pattern of the difference is plotted in figure 5.6. The largest difference between the $\Delta T_{2\text{-}4xCO2}$ and the TCR1 is in the region of the Arctic Ocean to the west of the Barents Sea to the East coast of Greenland. In the region surrounding the maximum, to the east of Greenland extending through the Barents Sea the TCR1 warms more than $\Delta T_{2\text{-}4xCO2}$. There is a 3.0 K difference in the North Atlantic Ocean off the coast of Maine and extends to the mid-North Atlantic Ocean. Throughout the midlatitudes and tropical regions, the difference is approximately 1.0 K. The Southern Ocean region has a very small difference between the TCR1 and $\Delta T_{2\text{-}4xCO2}$. Another local maximum is in the Weddell Sea to the east of the Antarctic peninsula, with a value of 2-3 K (Fig. 5.6). Thus, the pattern of surface warming over high latitude oceans is a key factor in controlling the rate of temperature increase with respect to each doubling of CO_2. The temperature in these regions is regulated by ocean currents, such as the Atlantic Meridional Overturning Current (AMOC). The fact that the TCR1 is displays more warming in parts of the Arctic ocean could be a consequence of changes in the AMOC circulation [Sévellec et al., 2017].

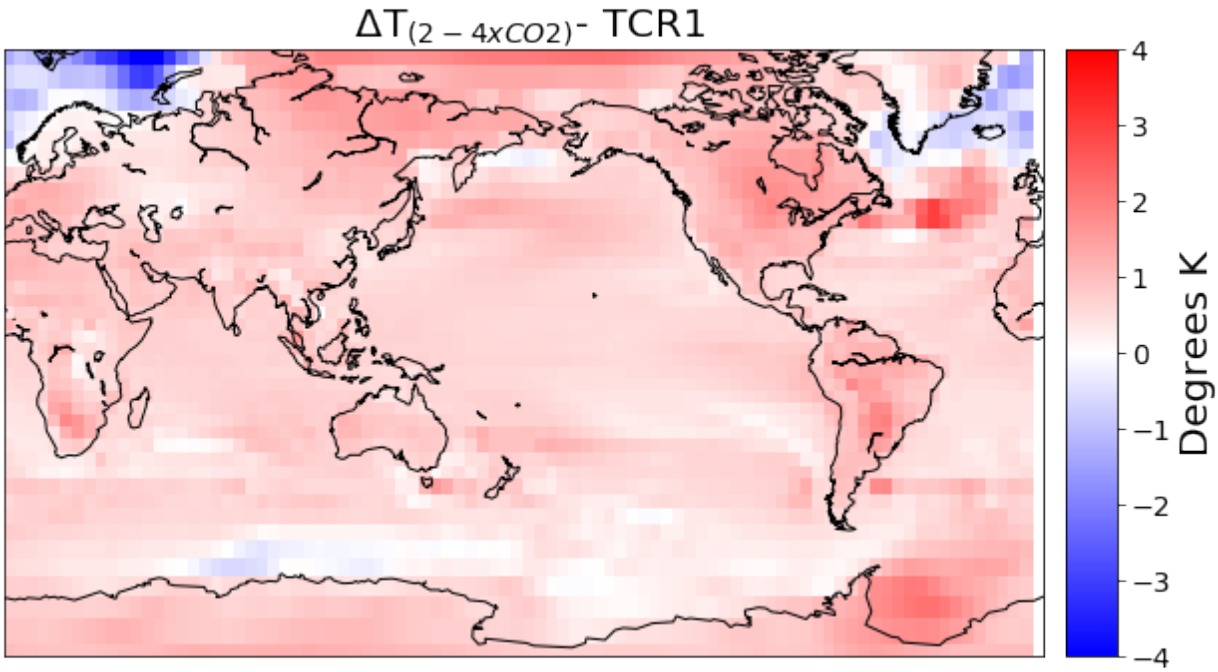

Figure 5.6: Map of the difference between $\Delta T_{2\text{-}4xCO2}$ and the TCR.

5.4 Correlation with temperature change

The total warming over the length of the 1%/year forcing scenario is calculated using the difference between the 1995-2005 average and the 1850 average, and referred to as ΔT. ΔT is correlated with TCR1, $\Delta T_{2\text{-}4xCO2}$, and the equilibrium climate sensitivity (ECS) from the large ensemble (Fig. 5.7). $\Delta T_{2\text{-}4xCO2}$ and the ECS have weak correlations with ΔT. TCR, on the other hand, explains about 78% of the variance observed in ΔT and is statistically significant at the 95% level. The three metrics are also correlated with the temperature change (ΔT) in the historical forcing scenario, and correlation coefficients for TCR1, ECS, and $\Delta T_{2\text{-}4xCO2}$ are less than 0.22 regardless of the subset of 68 of the 100 members that are selected.

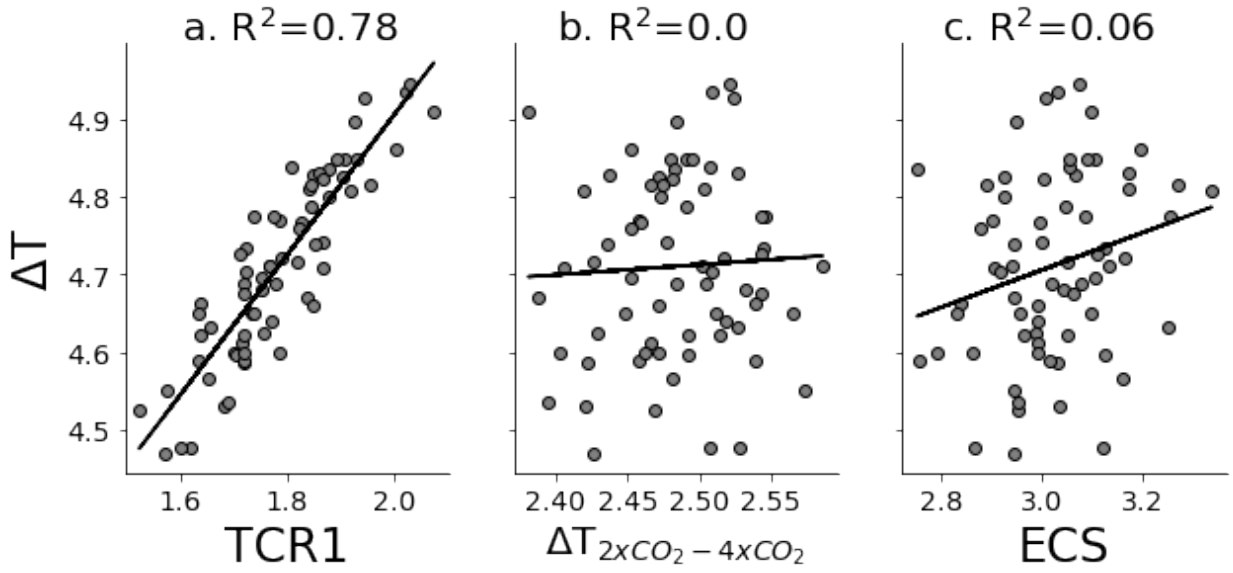

Figure 5.7: Correlation coefficient of a. TCR1, b. $\Delta T_{2\text{-}4xCO2}$ for the second doubling of CO2, and c.ECS vs. ΔT relative to 1850 in the 1% per year scenario. The linear regression line is plotted in black and R^2 values are listed.

5.4.1 Spatial pattern

Figure 5.8 shows the spatial pattern of the correlation coefficient resulting from the regression of ΔT at each latitude (same definition as in Fig. 5.7) with the TCR1 (blue line) and $\Delta T_{2\text{-}4xCO2}$ (red line) for each latitude. The TCR1 has a maximum correlation with ΔT at 70°S (sine of latitude 0.93, R^2= 0.9), and two secondary maxima of about 0.89 are located from in the tropical region and from 60°N- 80°N. The correlation is weakest in the regions of 55°S to 30°S with the R^2 ranging from 0.5 to 0.61. Thus, the TCR can generally explain the variance in ΔT in the high latitude and tropical regions and explains less in the Southern Ocean region (Fig. 5.8). Consistent with the global average (Fig. 5.7), the correlation coefficient for $\Delta T_{2\text{-}4xCO2}$ is smaller than the TCR1 for all latitudes. The maximum correlation of 0.1 occurs at about 35° S (sin(latitude) = -0.58) and $\Delta T_{2\text{-}4xCO2}$ explains less than 10% of the variability in ΔT elsewhere (Fig. 5.8). This shows that the TCR (vs. $\Delta T_{2\text{-}4xCO2}$) is the best predictor for warming (over the length of the 1% per year forcing runs) regardless of the region considered.

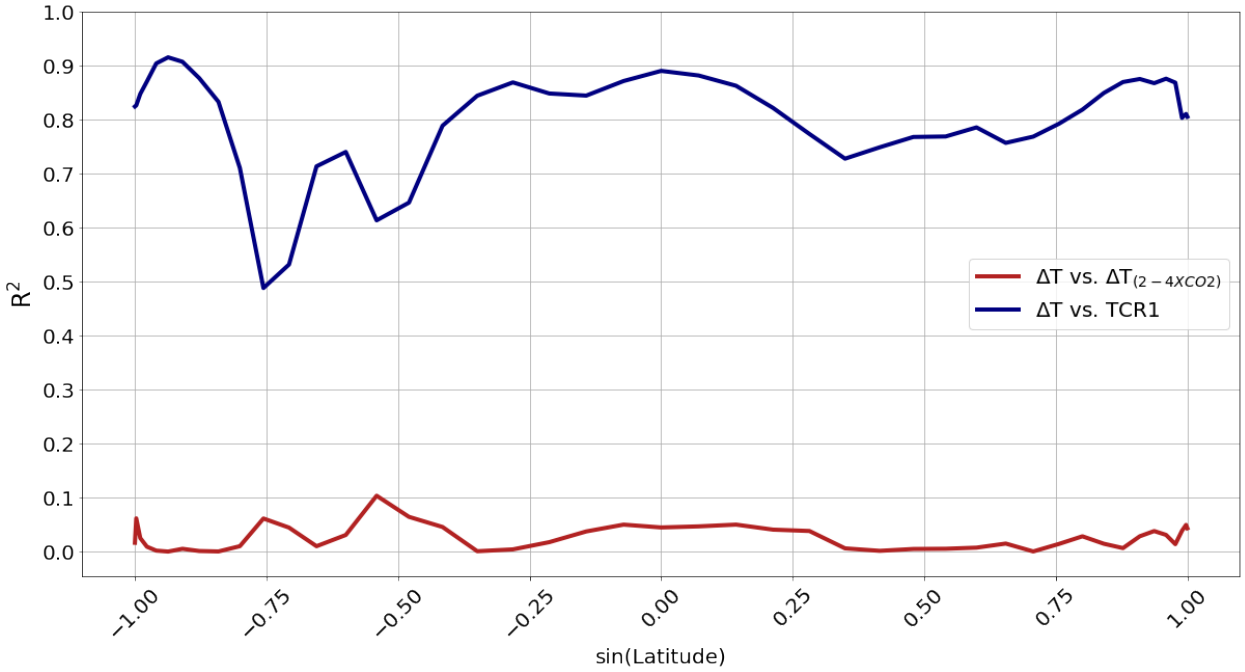

Figure 5.8: Correlation coefficient of TCR1 and ΔT for the second doubling of CO2 vs. ΔT relative to 1850 average in the 1% per year scenario.

5.4.2 Relationship of warming in the 20th century vs. 21st century

The large ensemble historical runs are extended to the 21st century under RCP4.5 and RCP8.5 forcing scenarios. ΔT for the RCP4.5 and RCP8.5 scenarios are calculated using the difference of the average surface temperature of the last decade of the run from the first decade. Figure 5.9a shows the correlation between the 21st century warming forced by RCP4.5 and the change in temperature in the 20th century. The correlation coefficient is equal to -0.0324, however the correlation is not statistically significant. There is also a slight negative correlation between the future warming forced by RCP8.5 (Fig. 5.9b), with an r-value of -0.067. This correlation is not statistically significant to the 95% level, with a p-value of 0.0507. The lack of correlation with future warming scenarios is evidence that the internal variability driving the variations in TCR have a characteristic time scale much shorter than a century.

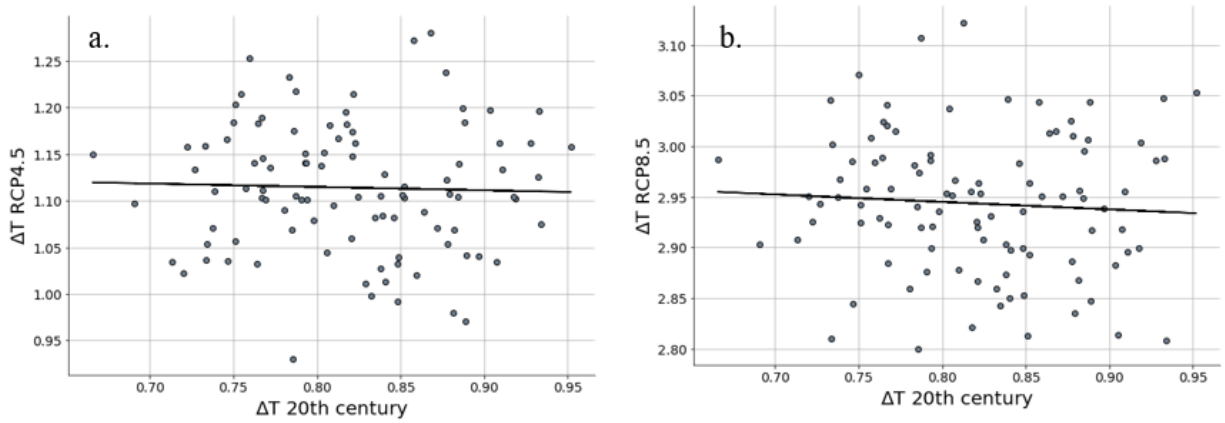

Figure 5.9: Correlation coefficient from the temperature change from 1850-2005 (relative to 1859-1882 average) and the temperature change in the 21st century (relative to 2006-2016 average). Panel a: forced by RCP 4.5 ensemble. Panel b: Forced by RCP8.5 scenario. Regression lines are overlaid.

5.5 Conclusion

The magnitude of transient warming due to the second doubling of CO_2 is greater than the warming from the first doubling (in 1% scenario). Thus, ΔT does not increase linearly in time (Fig. 5.1). This inconstancy is quantified by calculating the ratio of TCR1 to TCR2 for all 68 model members within the 1% per year increase in CO_2 scenario. The ratio of the two would equal 0.5 if the increase in temperature per doubling of CO_2 was constant. The ratio of the two is observed to vary from 0.38 to 0.47 in the large ensemble (Fig. 5.2). The ratio is minimized at about 70°N and maximized around 55°S. The largest variability in the ratio is located at 70°S, along the Antarctic edge (Fig. 5.4). The spatial pattern of the warming during the first and second half of the simulation is also investigated. Regions in the Arctic Ocean, North Atlantic Ocean, and the Southern Ocean are observed to have the greatest warming differences for the first and second halves of the simulation (Fig. 5.5 and 5.6). It is speculated that these spatial patterns are a result of ice-albedo feedback, ocean circulation, and heat content which could cause a delay in warming response.

Gregory et al. (2015) explains the possible reasons for the discrepancy between TCR1 and the TCR2. One factor is the inconstancy of ocean heat uptake rate and climate feedback parameter (λ). As time elapses in the 1% per year forcing run, the pattern and magnitude of OHC could evolve to approach the equilibrated temperatures causing a large shift in the temperature change for the second doubling vs. the first [Grose et al., 2018]. Another reason for the difference is that the forcing from CO_2 is assumed to increase logarithmically with concentration [Gregory et al., 2015]. In the large ensemble, the forcing due to CO_2 at time of doubling is F_{2xCO2}= 3.73 Wm^{-2} (centered average around year 70 and the same for all ensemble members, $2xF_{2xCO2}$ is 7.47 Wm^{-2}). The forcing due to a quadrupling of CO_2 is 7.83 Wm^{-2} in the large ensemble. Thus, the magnitude of forcing is 8.5% greater in the second half of the run in the large ensemble. This may lend some explanation to the lack of correlation between ΔT for the second doubling of CO_2 vs. the first. Estimates of the TCR from historical observations assume a constant F_{2xCO2} (Eq. 1.4). The forcing for the second doubling of CO_2 is observed to be larger than the first in the large ensemble. Since estimates of the TCR assume a constant value for F_{2xCO2} (Eq. 1.4) this inconsistency could add additional uncertainty to estimates of the TCR made from the historical record [Gregory et al., 2015].

6. CONCLUSION

Using a 100-member large ensemble experiment from the MPI Earth System Model (MPI-ESM1.1), we quantify the impact of internal variability on estimates of the TCR from the 20th century historical record. The TCR is a common climate science metric that quantifies the magnitude of short term warming in response to an increase in greenhouse gasses. The TCR is defined as the increase in temperature at the time of CO_2 doubling in a 1% per year increase scenario. Since the TCR occurs in an idealized forcing scenario it is common to estimate the TCR from the observational record [Gregory et al., 2015; Otto et al., 2013; Lewis and Curry, 2015].

6.1 The effect of internal variability on estimates of the TCR

We use a common method to estimate the TCR from historical observations (equation 1.4) and find that TCR estimates from individual members of the ensemble vary from 1.34 K to 1.92 K. Thus, the TCR estimated from any particular ensemble member may be strongly influenced by internal variability. Since observations of the climate system are limited to a single realization of internal variability, our results suggest caution in concluding that TCR estimated from the historical record reflects our climate system's actual TCR.

The majority of the variability in the TCR is located in the high latitude regions (Fig. 3.3), this is an artifact of polar amplification [Screen and Simmonds, 2010]. Via linear regression of ΔT_{region} and the TCR, it is found that the Northern Hemisphere explains a majority of the variability in the TCR, with some contribution from the tropical region (Fig. 3.5). OLS regression using ΔT from 4 regions in a 30° longitude x 15° latitude grid, can explain about 79% of the variability in the TCR (Fig. 3.6). In general, OLS regression using ΔT in the extra-tropical Northern hemisphere and the tropics (Fig. 3.7).

6.2 Controlling factors of the variability in the TCR

The Earth's linearized energy budget equation provides a framework with which the observed internal variability can be analyzed (Eq. 1.1). We find that TCR correlates most strongly with the

fraction of OHC that is stored below 100 meters (R^2 of the correlation is 0.56). This makes some sense since models that sequester more energy in the deep ocean would be expected to show less warming and therefore have a lower TCR. The climate feedback parameter is also found to explain 22% of the variance present in the TCR. Total accumulated energy, OHC total, and the OHC below 100 meters all have weak linear relationships with the TCR.

Using two variables fraction of energy stored below 100 m and energy radiated back to space provides an even better fit, explaining 71% of the variability in the TCR. This is consistent with energy conservation: the more energy that is stored in the deep ocean or radiated back to space, the less surface warming the Earth experiences.

6.3 Time dependence of the TCR

The true TCR of the large ensemble is investigated in a 68-member ensemble forced by 1% per year increase in CO_2 forcing scenario. It is found that the magnitude of transient warming due to the second doubling of CO_2 (year 140) is greater than the warming from the first doubling (year 70 in 1% scenario). The ratio of warming from the first to the second doubling of CO_2 varies from 0.38 to 0.47 in the large ensemble. The inconstancy of the TCR with time is another reason why estimates from the historical record should be interpreted with skepticism.

6.4 Concluding remarks

The TCR estimated from the 20th century record may not be a good estimate of our climate system's true value due to internal variability. We find here that internal variability is confounding our estimates of the TCR from the 20th century record. The variability observed can be traced back to variability in the processes that move energy into the deep ocean and in how much energy is radiated back to space. Estimates of the TCR made from the 20th century historical record represents one realization of the climate system and therefore may not characterize the true climate sensitivity.

REFERENCES

Allan, R. P., 2017: Global energy budget: Elusive origin of warming slowdown. *Nature Clim. Change*, **7 (5)**, 316–317, URL http://dx.doi.org/10.1038/nclimate3282.

Andrews, T., and M. J. Webb, 2018: The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming. *Journal of Climate*, **31 (2)**, 641–654, doi: 10.1175/JCLI-D-17-0087.1, URL http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0087.1.

Armour, K. C., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nature Clim. Change*, **7**, URL http://dx.doi.org/10.1038/nclimate3278.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2012: Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, **26 (13)**, 4518–4534, doi:10.1175/JCLI-D-12-00544.1, URL https://doi.org/10.1175/JCLI-D-12-00544.1.

Brown, P. T., W. Li, J. H. Jiang, and H. Su, 2016: Spread in the magnitude of climate model interdecadal global temperature variability traced to disagreements over high-latitude oceans. *Geophysical Research Letters*, **43 (24)**, 12,543–12,549, doi:10.1002/2016GL071442, URL http://dx.doi.org/10.1002/2016GL071442.

Cowtan, K., and R. G. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends: Coverage Bias in the HadCRUT4 Temperature Series. *Quarterly Journal of the Royal Meteorological Society*, **140 (683)**, 1935–1944, doi:10.1002/qj.2297, URL http://doi.wiley.com/10.1002/qj.2297.

Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, **38 (3)**, 527–546, doi:10.1007/s00382-010-0977-x, URL https://doi.org/10.1007/s00382-010-0977-x.

Dessler, A., and M. Zelinka, 2015: Climate and Climate Change| Climate Feedbacks. *Encyclopedia of Atmospheric Sciences*, Elsevier, 18–25, doi:10.1016/B978-0-12-382225-3.00471-0, URL http://linkinghub.elsevier.com/retrieve/pii/B9780123822253004710.

Dessler, A. E., T. Mauritsen, and B. Stevens, 2018: The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmospheric Chemistry and Physics Discussions*, 1–21, doi:10.5194/acp-2017-1236, URL https://www.atmos-chem-phys-discuss.net/acp-2017-1236/.

Forster, P. M., 2016: Inference of Climate Sensitivity from Analysis of Earth's Energy Budget. *Annual Review of Earth and Planetary Sciences*, **44 (1)**, 85–106, doi:10.1146/annurev-earth-060614-105156, URL https://doi.org/10.1146/annurev-earth-060614-105156.

Frankcombe, L. M., M. H. England, M. E. Mann, and B. A. Steinman, 2015: Separating Internal Variability from the Externally Forced Climate Response. *Journal of Climate*, **28 (20)**, 8184–8202, doi:10.1175/JCLI-D-15-0069.1, URL https://doi.org/10.1175/JCLI-D-15-0069.1.

Fyfe, J. C., N. P. Gillett, and F. W. Zwiers, 2013: Overestimated global warming over the past 20 years. *Nature Climate Change*, **3 (9)**, 767–769, doi:10.1038/nclimate1972, URL http://www.nature.com/doifinder/10.1038/nclimate1972.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5: Climate Changes in MPI-ESM. *Journal of Advances in Modeling Earth Systems*, **5 (3)**, 572–597, doi:10.1002/jame.20038, URL http://doi.wiley.com/10.1002/jame.20038.

Gregory, J. M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, **43 (8)**, 3911–3920, doi:10.1002/2016GL068406, URL http://dx.doi.org/10.1002/2016GL068406, https://andthentheresphysics.wordpress.com/2016/04/17/time-dependent-feedbacks/.

Gregory, J. M., T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response parameter under increasing CO2. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **373 (2054)**, doi:10.1098/rsta.2014.0417, URL http://rsta.royalsocietypublishing.org/content/373/2054/20140417.abstract.

Gregory, J. M., and P. M. Forster, 2008: Transient climate response estimated from radiative forcing and observed temperature change. *Journal of Geophysical Research: Atmospheres*, **113 (D23)**, n/a–n/a, doi:10.1029/2008JD010405, URL http://dx.doi.org/10.1029/2008JD010405.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *Journal of Climate*, **15 (22)**, 3117–3121, doi:10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2, URL https://doi.org/10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2.

Grose, M. R., J. Gregory, R. Colman, and T. Andrews, 2018: What climate sensitivity index is most useful for projections?: What climate sensitivity index is most useful for projections? *Geophysical Research Letters*, doi:10.1002/2017GL075742, URL http://doi.wiley.com/10.1002/2017GL075742.

Hawkins, E., and R. Sutton, 2009: The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, **90 (8)**, 1095–1108, doi:10.1175/2009BAMS2607.1, URL http://journals.ametsoc.org/doi/10.1175/2009BAMS2607.1.

Hedemann, C., T. Mauritsen, J. Jungclaus, and J. Marotzke, 2017: The subtle Origins of surface-warming hiatuses. *Nature Clim. Change*, **7**, 336–340, doi:10.1038/nclimate3274, URL http://www.nature.com/nclimate/journal/v7/n5/full/nclimate3274.html.

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Prein-

dustrial Forcing. *Journal of Climate*, **23 (9)**, 2418–2427, doi:10.1175/2009JCLI3466.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009JCLI3466.1.

Huber, M., U. Beyerle, and R. Knutti, 2014: Estimating climate sensitivity and future temperature in the presence of natural climate variability. *Geophysical Research Letters*, **41 (6)**, 2086–2092, doi:10.1002/2013GL058532, URL http://dx.doi.org/10.1002/2013GL058532.

Huber, M., and R. Knutti, 2012: Anthropogenic and natural warming inferred from changes in Earth/'s energy balance. *Nature Geosci*, **5 (1)**, 31–36, doi:10.1038/ngeo1327, URL http://dx.doi.org/10.1038/ngeo1327.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society*, **96 (8)**, 1333–1349, doi:10.1175/BAMS-D-13-00255.1, URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00255.1.

Lewis, N., and J. A. Curry, 2015: The implications for climate sensitivity of AR5 forcing and heat uptake estimates. *Climate Dynamics*, **45 (3)**, 1009–1023, doi:10.1007/s00382-014-2342-y, URL https://doi.org/10.1007/s00382-014-2342-y.

Marotzke, J., and P. M. Forster, 2015: Forcing, feedback and internal variability in global temperature trends. *Nature*, **517 (7536)**, 565–570, URL http://dx.doi.org/10.1038/nature14117.

Marvel, K., R. Pincus, G. A. Schmidt, and R. L. Miller, 2018: Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, doi:10.1002/2017GL076468, URL http://doi.wiley.com/10.1002/2017GL076468.

Mauritsen, T., and R. Pincus, 2017: Committed warming inferred from observations. *Nature Climate Change*, **7 (9)**, 652–655, doi:10.1038/nclimate3357, URL http://www.nature.com/doifinder/10.1038/nclimate3357.

Medhaug, I., M. B. Stolpe, E. M. Fischer, and R. Knutti, 2017: Reconciling controversies about the global warming hiatus. *Nature*, **545 (7652)**, 41–47, doi:10.1038/nature22315, URL http://www.nature.com/doifinder/10.1038/nature22315.

Murphy, D. M., S. Solomon, R. W. Portmann, K. H. Rosenlof, P. M. Forster, and T. Wong, 2009: An observationally based energy balance for the Earth since 1950. *Journal of Geophysical Research: Atmospheres*, **114 (D17)**, n/a–n/a, doi:10.1029/2009JD012105, URL http://dx.doi.org/10.1029/2009JD012105.

North, G. R., and K.-Y. Kim, 2017: *Energy Balance Climate Models*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, doi:10.1002/9783527698844, URL http://doi.wiley.com/10.1002/9783527698844.

Olson, R., R. Sriver, W. Chang, M. Haran, N. M. Urban, and K. Keller, 2013: What is the effect of unresolved internal climate variability on climate sensitivity estimates? *Journal of Geophysical Research: Atmospheres*, **118 (10)**, 4348–4358, doi:10.1002/jgrd.50390, URL http://dx.doi.org/10.1002/jgrd.50390.

Otto, A., and Coauthors, 2013: Energy budget constraints on climate response. *Nature Geosci*, **6 (6)**, 415–416, URL http://dx.doi.org/10.1038/ngeo1836.

Padilla, L. E., G. K. Vallis, and C. W. Rowley, 2011: Probabilistic Estimates of Transient Climate Sensitivity Subject to Uncertainty in Forcing and Natural Variability. *Journal of Climate*, **24 (21)**, 5521–5537, doi:10.1175/2011JCLI3989.1, URL http://dx.doi.org/10.1175/2011JCLI3989.1.

Raper, S. C. B., J. M. Gregory, and R. J. Stouffer, 2002: The Role of Climate Sensitivity and Ocean Heat Uptake on AOGCM Transient Temperature Response. *Journal of Climate*, **15 (1)**, 124–130, doi:10.1175/1520-0442(2002)015<0124:TROCSA>2.0.CO;2, URL https://doi.org/10.1175/1520-0442(2002)015<0124:TROCSA>2.0.CO;2.

Richardson, M., K. Cowtan, E. Hawkins, and M. B. Stolpe, 2016: Reconciled climate response

estimates from climate models and the energy budget of Earth. *Nature Clim. Change*, **6 (10)**, 931–935, URL http://dx.doi.org/10.1038/nclimate3066.

Screen, J. A., and I. Simmonds, 2010: The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, **464 (7293)**, 1334–1337, doi:10.1038/nature09051, URL http://www.nature.com/articles/nature09051.

Skeie, R. B., T. Berntsen, M. Aldrin, M. Holden, and G. Myhre, 2014: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. *Earth System Dynamics*, **5 (1)**, 139–175, doi:10.5194/esd-5-139-2014, URL http://www.earth-syst-dynam.net/5/139/2014/.

Stott, P., P. Good, G. Jones, N. Gillett, and E. Hawkins, 2013: The upper end of climate model temperature projections is inconsistent with past warming. *Environmental Research Letters*, **8 (1)**, 014 024, doi:10.1088/1748-9326/8/1/014024, URL http://stacks.iop.org/1748-9326/8/i=1/a=014024?key=crossref.945b685403738a300549a5978df98ca0.

Sévellec, F., A. V. Fedorov, and W. Liu, 2017: Arctic sea-ice decline weakens the Atlantic Meridional Overturning Circulation. *Nature Climate Change*, **7 (8)**, 604–610, doi:10.1038/nclimate3353, URL http://www.nature.com/doifinder/10.1038/nclimate3353.

Von Schuckmann, K., and Coauthors, 2016: An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, **6 (2)**, 138–144, doi:10.1038/nclimate2876, URL http://www.nature.com/doifinder/10.1038/nclimate2876.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earths energy budget. *Nature Geoscience*, **9 (12)**, 871–874, doi:10.1038/ngeo2828, URL http://www.nature.com/articles/ngeo2828.